

1/14/19

From: Fredman, Jeffrey
Sent: Wednesday, February 11, 2004 11:07 AM
To: STIC-Biotech/ChemLib
Cc: Chernyshev, Olga
Subject: FW: sequence search request

PLEASE RUSH.

I Approve this large search.

Jeff Fredman



-----Original Message-----

From: Chernyshev, Olga
Sent: Wednesday, February 11, 2004 10:12 AM
To: Fredman, Jeffrey
Subject: sequence search request

Please authorize the following sequence search request:

Please search US case 09/901,187 SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 in regular and pending databases. Only direct hits, because they claim these short peptides precisely, no open language.

Thank you very much!

Olga N. Chernyshev
AU1646
REM 4E84
20870
mail box 4D70



RECEIVED
FEB 11 2004
STIC-BIOTECH/CHEM LIBRARY
(STIC)

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: *2/11/04*
Date Completed: *2/12/04*
Searcher Prep/Review: *2/11/04*
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: *12*
AA Sequences: *12*
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: *OSP*
Sequence Sys.: *OSP*
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
 (without alignments)
 59.419 Million cell updates/sec

Title: US-09-901-187C-11
 Perfect score: 44
 Sequence: 1 THHHPSP 7

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues
 Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

16 35 79.5 570 12 US-10-369-493-3132
 17 35 79.5 612 12 US-10-369-493-9134
 18 35 79.5 1821 12 US-10-372-683-22
 19 34 77.3 58 15 US-10-083-357-1179
 20 34 77.3 253 15 US-10-102-627-66
 21 34 77.3 338 10 US-09-847-010-6
 22 34 77.3 338 15 US-10-235-939-6
 23 34 77.3 344 9 US-09-012-135A-3
 24 34 77.3 344 12 US-10-295-027-203
 25 34 77.3 344 12 US-10-173-999-101
 26 34 77.3 344 15 US-10-059-585-34
 27 34 77.3 344 15 US-10-171-311-214
 28 34 77.3 347 10 US-09-974-298-136
 29 34 77.3 354 12 US-10-108-260A-3493
 30 34 77.3 417 12 US-10-369-493-17203
 31 34 77.3 439 9 US-09-741-669-300
 32 34 77.3 439 9 US-09-815-242-10284
 33 34 77.3 461 9 US-09-731-557A-2
 34 34 77.3 802 11 US-09-946-374-260
 35 34 77.3 802 12 US-10-015-387A-260
 36 34 77.3 802 12 US-10-006-130A-260
 37 34 77.3 802 12 US-10-199-672-312
 38 34 77.3 802 12 US-10-006-172A-260
 39 34 77.3 802 12 US-10-187-749-312
 40 34 77.3 802 12 US-10-194-457-312
 41 34 77.3 802 12 US-10-184-642-312
 42 34 77.3 802 12 US-10-196-747-312
 43 34 77.3 802 12 US-10-015-392A-260
 44 34 77.3 802 12 US-10-017-253A-260
 45 34 77.3 802 12 US-10-173-689-312

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

卷之三

Publication No. US20030124679A1

GENERAL INFORMATION:
 / APPLICANT: Short, Jay M.
 / APPLICANT: Paradkar, Ashish
 / APPLICANT: Varoglu, Mustafa
 / APPLICANT: Mathur, Eric J.
 / TITLE OF INVENTION: GLYCOSYLATED KINAMYCINS AND METHODS OF
 / FILE REFERENCE: 09010-280001
 / CURRENT APPLICATION NUMBER: US/10/187,267A
 / CURRENT FILING DATE: 2003-01-27
 / PRIOR APPLICATION NUMBER: US 60/301,401
 / PRIOR FILING DATE: 2001-06-27
 / NUMBER OF SEQ ID NOS: 65
 / SEQ ID NO: 19
 / LENGTH: 148
 / TYPE: PRT
 / ORGANISM: Streptomyces murayamaensis ATCC 21414
 / US-10-187-267A-19

Query Match 97.7%; Score 43; DB 15; Length 148;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THIIEHPS 7
 Db 93 THVIEHPS 99

RESULT 3
 US-09-801-944B-233

Sequence 233, Application US/09801944B
 / Publication No. US20040014169A1

GENERAL INFORMATION:
 / APPLICANT: Vogeli, Gabriel

/ APPLICANT: Wood, Linda S.

TITLE OF INVENTION: No. US20040014169A1 e G Protein-Coupled Receptors

FILE REFERENCE: 00100US1

CURRENT APPLICATION NUMBER: US/09/801,944B
 CURRENT FILING DATE: 2003-08-01
 PRIOR APPLICATION NUMBER: 60/187,828
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,715
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,929
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,930
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,825
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,833
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,830
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,829
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,582
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,581
 PRIOR FILING DATE: 2000-03-08
 Remaining Prior Application data removed - See File Wrapper or PAlM.
 NUMBER OF SEQ ID NOS: 273

SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 233
 LENGTH: 178
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-801-944B-233

Query Match 88.6%; Score 39; DB 12; Length 178;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 NUMBER OF SEQ ID NOS: 109

Query Match 84.1%; Score 37; DB 15; Length 262;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 NUMBER OF SEQ ID NOS: 238

Qy 2 THIIEHPS 7
 Db 102 THHIEHPS 107

RESULT 4
 US-10-341-434-14
 / Sequence 14, Application US/10341434
 / Publication No. US20030215835A1

GENERAL INFORMATION:

APPLICANT: Origene Technologies
 / TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
 / FILE REFERENCE: 9U 204 205 RI
 / CURRENT APPLICATION NUMBER: US/10/341,434
 / CURRENT FILING DATE: 2003-07-18
 / PRIOR APPLICATION NUMBER: US 60/348,164
 / PRIOR FILING DATE: 2002-01-15
 / PRIOR APPLICATION NUMBER: US 60/348,119
 / PRIOR FILING DATE: 2002-01-15
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO: 14
 / LENGTH: 402
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-10-341-434-14

Query Match 88.6%; Score 39; DB 12; Length 402;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIIEHPS 7
 Db 247 THCHEHPS 253

RESULT 5
 US-10-183-116-99

/ Sequence 99, Application US/10183116
 / Publication No. US20030092035A1

GENERAL INFORMATION:

APPLICANT: Anderson, David J.

APPLICANT: Dong, Xinzhang

APPLICANT: Zylka, Mark

APPLICANT: Simon, Melvin

APPLICANT: Han, Sang-kyou

TITLE OF INVENTION: PAIN SIGNALING MOLECULES

FILE REFERENCE: CALTE.4C1CP1

CURRENT APPLICATION NUMBER: US/10/183,116

CURRENT FILING DATE: 2002-06-26

PRIOR APPLICATION NUMBER: US 60/222,344

PRIOR FILING DATE: 2000-08-01

PRIOR APPLICATION NUMBER: US 60/202,027

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 09/704,707

PRIOR FILING DATE: 2000-11-03

PRIOR APPLICATION NUMBER: US 60/285,493

PRIOR FILING DATE: 2001-04-19

PRIOR APPLICATION NUMBER: US 09/849,869

PRIOR FILING DATE: 2001-05-04

NUMBER OF SEQ ID NOS: 109

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 99

LENGTH: 262

TYPE: PRT

ORGANISM: Mus musculus

US-10-183-116-99

Qy 2 HIHHPs 7
 Db 208 HLHHPs 213

RESULT 6
 US-10-183-116-91
 ; Sequence 91, Application US/10183116
 ; Publication No. US20030092035A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David J.
 ; FILE REFERENCE: CALTE.4C1CP1
 ; CURRENT APPLICATION NUMBER: US/10/183,116
 ; CURRENT FILING DATE: 2002-06-26
 ; PRIOR APPLICATION NUMBER: US 60/222,344
 ; PRIOR FILING DATE: 2000-08-01
 ; PRIOR APPLICATION NUMBER: US 60/202,027
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 09/704,707
 ; PRIOR FILING DATE: 2000-11-03
 ; PRIOR APPLICATION NUMBER: US 60/285,493
 ; PRIOR FILING DATE: 2001-04-19
 ; PRIOR APPLICATION NUMBER: US 09/849,869
 ; PRIOR FILING DATE: 2001-05-04
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 91
 ; LENGTH: 270
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

Query Match 84.1%; Score 37; DB 15; Length 270;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHPs 7
 Db 216 HLHHPs 221

RESULT 7
 US-10-115-482-50
 ; Sequence 50, Application US/10115482
 ; Publication No. US20030212257A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Spytel, et al.
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
 ; TITLE OF INVENTION: AND METHODS
 ; TITLE OF INVENTION: OF USING THE SAME
 ; FILE REFERENCE: 21404-322D
 ; CURRENT APPLICATION NUMBER: US/10/115,482
 ; CURRENT FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: 60/281,086
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: 60/281,136
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: 60/281,863
 ; PRIOR FILING DATE: 2001-04-05
 ; PRIOR APPLICATION NUMBER: 60/281,906
 ; PRIOR FILING DATE: 2001-04-10
 ; PRIOR APPLICATION NUMBER: 60/283,512
 ; PRIOR FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: 60/285,325
 ; PRIOR FILING DATE: 2001-04-19
 ; PRIOR APPLICATION NUMBER: 60/285,890
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 60/286,068
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: 60/286,292
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 60/287,213
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/288,257
 ; PRIOR FILING DATE: 2001-05-02

Query Match 84.1%; Score 37; DB 12; Length 911;
 Best Local Similarity 83.3%; Pred. No. 5.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHPs 7
 Db 320 HLHHPs 325

RESULT 8
 US-10-115-482-48
 ; Sequence 48, Application US/10115482
 ; Publication No. US20030212257A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Spytel, et al.
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
 ; TITLE OF INVENTION: AND METHODS
 ; TITLE OF INVENTION: OF USING THE SAME
 ; FILE REFERENCE: 21404-322D
 ; CURRENT APPLICATION NUMBER: US/10/115,482
 ; CURRENT FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: 60/281,086
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: 60/281,136
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: 60/281,863
 ; PRIOR FILING DATE: 2001-04-05
 ; PRIOR APPLICATION NUMBER: 60/281,906
 ; PRIOR FILING DATE: 2001-04-10
 ; PRIOR APPLICATION NUMBER: 60/283,512
 ; PRIOR FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: 60/285,325
 ; PRIOR FILING DATE: 2001-04-19
 ; PRIOR APPLICATION NUMBER: 60/285,890
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 60/286,068
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: 60/286,292
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 60/287,213
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/288,257
 ; PRIOR FILING DATE: 2001-05-02

PRIOR APPLICATION NUMBER: 60/291,134
 PRIOR FILING DATE: 2001-05-15
 PRIOR APPLICATION NUMBER: 60/282,020
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/291,725
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/294,771
 PRIOR FILING DATE: 2001-05-31
 NUMBER OF SEQ ID NOS: 149
 SEQ ID NO 48
 LENGTH: 915
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-115-482-48

Query Match 84.1%; Score 37; DB 12; Length 915;
 Best Local Similarity 83.3%; Pred. No. 5.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HIHHPs 7
 Db 320 HLHHPs 325

RESULT 9
 US-10-017-161-740
 Sequence 740, Application US/10017161
 Publication No. US20030143668A1
 GENERAL INFORMATION:
 APPLICANT: SUWA, MAKIKO
 APPLICANT: ASAI, KIYOSHI
 APPLICANT: AKIYAMA, YUTAKA
 APPLICANT: ABURATANI, HIROYUKI
 TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 FILE REFERENCE: 084335/0152
 CURRENT APPLICATION NUMBER: US/10/017,161
 CURRENT FILING DATE: 2002-12-18
 PRIOR APPLICATION NUMBER: JP 2001/246789
 PRIOR FILING DATE: 2001-06-18
 NUMBER OF SEQ ID NOS: 2430
 SEQ ID NO 740
 LENGTH: 1021
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-017-161-740

Query Match 84.1%; Score 37; DB 12; Length 1021;
 Best Local Similarity 83.3%; Pred. No. 6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HIHHPs 7
 Db 361 HLHHPs 366

RESULT 10
 US-10-335-687A-2

Sequence 2, Application US/10335687A
 Publication No. US20030166222A1
 GENERAL INFORMATION:
 APPLICANT: Meyers, Rachel E.
 APPLICANT: Millennium Pharmaceuticals, Inc.
 TITLE OF INVENTION: Uses Therefor
 FILE REFERENCE: MPI02-001P1RNM
 CURRENT APPLICATION NUMBER: US/10/335,687A
 CURRENT FILING DATE: 2003-01-02
 PRIOR APPLICATION NUMBER: 60/345,773
 PRIOR FILING DATE: 2002-01-02
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 5
 LENGTH: 1824
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-335-687A-5

RESULT 11
 US-10-335-687A-5
 Sequence 5, Application US/10335687A
 Publication No. US20030166222A1
 GENERAL INFORMATION:
 APPLICANT: Meyers, Rachel E.
 APPLICANT: Millennium Pharmaceuticals, Inc.
 TITLE OF INVENTION: Human Kinase Family Members and
 TITLE OF INVENTION: Uses Therefor
 FILE REFERENCE: MPI02-001P1RNM
 CURRENT APPLICATION NUMBER: US/10/335,687A
 CURRENT FILING DATE: 2003-01-02
 PRIOR APPLICATION NUMBER: 60/345,773
 PRIOR FILING DATE: 2002-01-02
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 5
 LENGTH: 1824
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-335-687A-5

Query Match 84.1%; Score 37; DB 12; Length 1824;
 Best Local Similarity 83.3%; Pred. No. 1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HIHHPs 7
 Db 1223 HLHHPs 1228

RESULT 12
 US-10-082-830-180
 Sequence 180, Application US/10082830
 Publication No. US20030077604A1
 GENERAL INFORMATION:
 APPLICANT: Sun, Yongming
 APPLICANT: Recipon, Herve
 APPLICANT: Salceda, Susana
 APPLICANT: Liu, Chenghua
 APPLICANT: Turner, Leah
 TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
 Genes and Proteins
 FILE REFERENCE: DEX-0249
 CURRENT APPLICATION NUMBER: US/10/082,830
 CURRENT FILING DATE: 2001-10-29
 PRIOR APPLICATION NUMBER: 60/243,802
 PRIOR FILING DATE: 2000-10-27
 NUMBER OF SEQ ID NOS: 282
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 180
 LENGTH: 46
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-082-830-180

Query Match 81.8%; Score 36; DB 15; Length 46;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THHHP 6
 Db 27 SHHHP 32

RESULT 13
 US-09-864-761-38417
 ; Sequence 38417, Application US/09864761
 ; Patent No. US20020048763A1

GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
 SEQ ID NO 38417
 LENGTH: 49
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:

OTHER INFORMATION: MAP TO Z98884.11
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.67
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
 OTHER INFORMATION: EST HUMAN HIT: BEB43827.1, EVALUATE 1.00e-24
 OTHER INFORMATION: SWISSPROT HIT: P56645, EVALUATE 7.00e-24
 US-09-864-761-38417

Query Match 79.5%; Score 35; DB 9; Length 49;
 Best Local Similarity 66.7%; Pred. No. 73;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIIHPS 7
 Db 22 HVHHPA 27

RESULT 14
 US-09-272-975-52
 ; Sequence 52, Application US/09272975
 ; Publication No. US20030027774A1

GENERAL INFORMATION:
 ; APPLICANT: Hendrickson, Ronald C.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Houghton, Raymond L.
 ; TITLE OF INVENTION: TUBERCULOSIS ANTIGENS AND METHODS
 ; TITLE OF INVENTION: OF USE THEREFOR
 ; FILE REFERENCE: 210121.474
 ; CURRENT APPLICATION NUMBER: US/09/272,975
 ; CURRENT FILING DATE: 1999-03-18
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 52
 ; LENGTH: 51
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 US-09-272-975-52

Query Match 79.5%; Score 35; DB 11; Length 51;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIIHPS 6
 Db 5 HVHHP 9

RESULT 15
 US-10-369-493-4227
 ; Sequence 4227, Application US/10369493
 ; Publication No. US20030233675A1

GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10-(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 4227
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 US-10-369-493-4227

Query Match 79.5%; Score 35; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIIHPS 6

Thu Feb 12 09:07:31 2004

Db 24 HHHHP 28

Search completed: February 11, 2004, 17:54:13
Job time : 24.6667 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
 Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
 (without alignments)
 72.127 Million cell updates/sec

Title: US-09-901-187C-11
 Perfect score: 44
 Sequence: 1 THIHHP 7
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 2833308

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	39	88.6	312	2	D86488		bacterial IS-élément IS10 transposase
2	39	88.6	402	2	I67760		- Esch transposase - Esch probable transposa
3	39	88.6	402	2	AE0737		hypothetical prote
4	37	84.1	66	2	C82702		hypothetical prote
5	37	84.1	69	2	C81075		hypothetical prote
6	37	84.1	611	2	T22855		ventral nervous sy
7	37	84.1	722	2	S57246		hypothetical prote
8	36	81.8	156	2	GB1900		hypothetical prote
9	36	81.8	376	2	S45107		phage virion prote
10	36	81.8	443	2	B81122		zinc finger trans
11	36	81.8	827	2	T39608		hypothetical prote
12	35	79.5	169	2	A82736		pyrroloquinoline Q
13	35	79.5	250	2	C83397		myocyte nuclear fa
14	35	79.5	251	2	S20455		73.8K alpha trans-
15	35	79.5	617	2	A56051		band 3-related pro
16	35	79.5	693	1	TNBFF7		band 3-related pro
17	35	79.5	1234	2	A34911		band 3-related pro
18	35	79.5	1237	2	A31789		band 3-related pro
19	35	79.5	1237	2	A56764		anion exchange pro
20	35	79.5	1240	2	S21086		brefeldin a resist
21	35	79.5	1530	2	S52239		hypothetical prote
22	35	79.5	1530	2	T52010		latent transformin
23	35	79.5	1820	2	A55494		hypothetical prote
24	34	77.3	108	2	AD1986		hypothetical prote
25	34	77.3	338	2	B64148		rubredoxin-NAD+ re
26	34	77.3	392	2	I39521		homoserine dehydro
27	34	77.3	417	2	E83806		Guanine deaminase
28	34	77.3	434	2	AE3448		Guanine deaminase
29	34	77.3	439	2	C65072		

ALIGNMENTS

30	34	77.3	439	2	D91098		hypothetical prote	
31	34	77.3	439	2	H85943		carbon catabolite	
32	34	77.3	474	2	S66480		hypothetical prote	
33	34	77.3	515	2	H86202		hypothetical prote	
	34	77.3	520	2	T30808		BRcore-NS-Z3 prote	
	34	77.3	566	2	A84485		hypothetical prote	
	36	34	77.3	704	2	S21911		hepatocyte nuclear
	37	34	77.3	774	2	JC6955		reverse transcript
	38	34	77.3	819	2	T10355		aryl hydrocarbon r
	39	34	77.3	916	2	JT0396		GTPase-activating
	40	34	77.3	920	2	JC7313		hypothetical prote
	41	34	77.3	1165	2	S27809		protein ECL - rat
	42	33	75.0	63	2	D82841		hypothetical prote
	43	33	75.0	65	2	T18064		
	44	33	75.0	75	2	S25719		
	45	33	75.0	250	2	T48684		

RESULT 1

D86488	bacterial IS-élément [Imported] - Arabidopsis thaliana (mouse-ear Cress)	C;Species: Arabidopsis thaliana	C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002	C;Accession: D86488			
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Raul, S.; White, O.; Alonso-Chin, C.W.; Chung, C.W.; Marziali, N.F.; Hughes, B.; Huizar, L.	Nature 408, 816-820, 2000	A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	A;Cross-references: GB:AE005172; NID:G11094742; PIDN:AAG29676.1; GSPDB:GN00141	A;Cross-references: GB:AE005172	A;Residues: 1-312 <STO>	C;Genetics:	C;Superfamily: Escherichia coli insertion sequence IS10 transposase
		A;Map position: 1	C;Map position: 1				
		C;Status: preliminary	C;Status: preliminary				
		A;Molecule type: DNA	A;Molecule type: DNA				

RESULT 2

I67760	transposase - Escherichia coli insertion sequence IS10	C;Species: Escherichia coli	C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 04-Mar-2000	C;Accession: I67760			
R;Bogosian, G.; Bilyeu, K.; O'Neil, J.P.	Gene 133, 17-22, 1993	A;Title: Genome rearrangements by residual IS10 elements in strains of Escherichia coli	A;Reference number: 153672; PMID:94040791	A;Accession: I67760			
		C;Status: preliminary; translated from GB/EMBL/DDBJ	A;Molecule type: DNA				
			A;Residues: 1-402 <RES>				
			A;Cross-references: GB:S67119; NID:9455674; PIDN:AAB28848.1; PID:9455676				
			A;Experimental source: plasmid pXT107; insertion sequence IS10				
			C;Genetics:				

A;Mobile element: insertion sequence IS10
 C;Superfamily: Escherichia coli insertion sequence IS10 transposase
 C;Keywords: DNA binding

Query Match 88.6%; Score 39; DB 2; Length 402;
 Best Local Similarity 85.7%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 THHHPS 7
 Db 247 THCHHPS 253

RESULT 3

AE0737
 probable transposase STY2056 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A;Note: this species has also been called *Salmonella typhi*
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AE0737
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.S.; Moule, S.; O'Gara, P.
Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi [imported] - *Salmonella enterica* serovar Typhi
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AE0737
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-402 <PAR>
 A;Cross-references: GB:AU1513382; PIDN:CAD05600.1; PID:g16503097; GSPDB:GN00176
 C;Genetics:
 A;Gene: STY2056
 C;Superfamily: Escherichia coli insertion sequence IS10 transposase

Query Match 88.6%; Score 39; DB 2; Length 402;
 Best Local Similarity 85.7%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 THHHPS 7
 Db 247 THCHHPS 253

Query Match 84.1%; Score 37; DB 2; Length 66;
 Best Local Similarity 85.7%; Pred. No. 3.1;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 THHHPS 7
 Db 41 THDHHPS 47

RESULT 4

C82702
 hypothetical protein XFL277 [imported] - *Xylella fastidiosa* (strain 9a5C)
 C;Species: *Xylella fastidiosa*
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C;Accession: C82702
 R;Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
 A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A;Reference number: A82515; MUID:20365717; PMID:10910347
 A;Note: for a complete list of authors see reference number A59328 below
 A;Accession: C82702
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-66 <SIM>
 A;Cross-references: GB:AB003961; GB:AE003849; PIDN:99106254; PIDN:AAF84086.1; GSPDB:GN001
 A;Experiment source: Strain 9a5C
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Aencio, M.; Alvaranga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kempfer, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Lai
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A;Authors: Martins, E.M.F.; Matsuoka, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 Rodriguez, F.G.; Nunes, L.R.; Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 A;Authors: Nunes, L.R.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tsuhako, M.H.; Vallada, H.; van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
 A;Reference number: A59328
 A;Contents: annotation
 C;Genetics:
 A;Gene: XFL277

Query Match 84.1%; Score 37; DB 2; Length 66;
 Best Local Similarity 85.7%; Pred. No. 3.1;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 THHHPS 7
 Db 41 THDHHPS 47

RESULT 5

C81075
 hypothetical protein NMB1503 [imported] - *Neisseria meningitidis* (strain MC58 serogroup C;Species: *Neisseria meningitidis*
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C;Accession: C81075
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
 A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58
 A;Reference number: AB1000; MUID:20175755; PMID:10710307
 A;Accession: CB1075
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-69 <TET>
 A;Cross-references: GB:AE002500; GB:AE002098; PIDN:97226744; PIDN:AAF41859.1; PID:972267
 A;Experimental source: serogroup B, strain MC58
 C;Genetics:
 A;Gene: NMB1503

Query Match 84.1%; Score 37; DB 2; Length 69;
 Best Local Similarity 83.3%; Pred. No. 3.3;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 THHHPS 7
 Db 40 HLHHHPS 45

RESULT 6

T22855
 hypothetical protein F57F5.4 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T22855
 R;Harris, B.
 Submitted to the EMBL Data Library, July 1996
 A;Reference number: Z19627
 A;Accession: T22855
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residues: 1-611 <WIL>
 A;Cross-references: EMBL:Z75953; PIDN:CA00100.1; GSPDB:GN00023; CESP:F57F5.4
 A;Experimental source: clone F57F5
 C;Genetics:
 A;Gene: CESP:F57F5.4
 A;Map position: 5
 A;Introns: 26/3; 56/2; 101/1; 150/3; 296/3; 382/3; 486/3; 570/1
 Query Match 84.1%; Score 37; DB 2; Length 611;
 Best Local Similarity 83.3%; Pred. No. 3.8;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 THHHPS 7
 Db 451 HLHHHPS 456

RESULT 7
 S57246 ventral nervous system defective protein - fruit fly (Drosophila melanogaster)
 N; Alternate names: NK-2 homeotic protein
 C; Species: Drosophila melanogaster
 C; Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 15-Oct-1999
 C; Accession: S57246; B33976
 R; Jimenez, F.; Martin-Morris, L.E.; Velasco, L.; Chu, H.; Sierra, J.; Rosen, D.R.; White EMBO J. 14, 3487-3495, 1995
 A; Title: vnd, a gene required for early neurogenesis of Drosophila, encodes a homeodomain
 A; Reference number: S57246; MUID:95354667; PMID:7628450
 A; Accession: S57246
 A; Molecule type: mRNA
 A; Residues: 1-722 <JIM>
 A; Cross-references: EMBL:X87141; NID:91045047; PIDN:CAA60619.1; PID:91045048
 R; Kim, Y.; Nirenberg, M.
 Proc. Natl. Acad. Sci. U.S.A. 86, 7716-7720, 1989
 A; Title: Drosophila NK-homeobox genes.
 A; Reference number: A33976; MUID:90046666; PMID:2573058
 A; Accession: B33976
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Residues: 476-631, 'VG' <KIM>
 A; Cross-references: GB:M27290; NID:9157635; PIDN:AAA28617.1; PID:9157636
 C; Genetics:
 A; Gene: FlyBase:rnd
 A; Cross-references: FlyBase:FBgn0003986
 C; Superfamily: unassigned homeobox proteins; homeobox homology
 C; Keywords: DNA binding; homeobox; nucleus; transcription regulation
 P; 545-601/Domain: homeobox homology <HOX>
 Query Match 84.1%; Score 37; DB 2; Length 722;
 Best Local Similarity 85.7%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 THHHPS 7
 Db 620 THPHHPS 626

RESULT 8
 GB1900 hypothetical protein NMA1319 [imported] - Neisseria meningitidis (strain Z2491 serogroup C; Species: Neisseria meningitidis
 C; Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C; Accession: G81900
 R; Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Kilee, S.R.; Morel, J.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, NATURE 404, 502-506, 2000
 A; Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A; Reference number: A81775; MUID:20222556; PMID:10761919
 A; Accession: G81900
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Residues: 1-156 <PAR>
 A; Cross-references: GB:AL162755; GB:AL157959; NID:97379742; PIDN:CAB64570.1; PID:9737999
 A; Experimental source: serogroup A, strain Z2491
 C; Genetics:
 A; Gene: NMA1318; NMA1319
 Query Match 81.8%; Score 36; DB 2; Length 156;
 Best Local Similarity 83.3%; Pred. No. 12;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HIHHP 7
 Db 138 HIHHPA 143

RESULT 9
 S45107

hypothetical protein 1 - *Erwinia carotovora*
 C; Species: *Erwinia carotovora*
 C; Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
 C; Accession: S45107
 R; Golby, P.; Jones, S.E.; Stephens, S.; Reeves, P.J.; Bcroft, B.; Stewart, G.; William submitted to the EMBL Data Library, May 1994
 A; Description: Global regulation of *Erwinia carotovora* exoenzyme virulence factors: mul
 A; Reference number: S45107
 A; Accession: S45107
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Cross-references: EMBL:X79474; NID:9496597; PIDN:CAA55982.1; PID:9496598
 A; Residues: 1-376 <COL>
 Query Match 81.8%; Score 36; DB 2; Length 376;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 THHHPS 6
 Db 283 SHHHP 288

RESULT 10
 B81122 phage virion protein, probable NMB1109 [imported] - *Neisseria meningitidis* (strain MC58
 C; Species: *Neisseria meningitidis*
 C; Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C; Accession: B81122
 R; Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J. Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Qi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000
 A; Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V A; Title: Complete Genome Sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A; Reference number: A81000; MUID:20175755; PMID:10710307
 A; Accession: B81122
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Cross-references: GB:AE002460; GB:AE002098; NID:97226335; PIDN:AAF41500.1; PID:972263
 A; Experimental source: serogroup B, strain MC58
 C; Genetics:
 A; Gene: NMB1109

Query Match 81.8%; Score 36; DB 2; Length 443;
 Best Local Similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 7
 Db 425 HIHHPA 430

RESULT 11
 T39608 zinc finger transcription factor - fission yeast (*Schizosaccharomyces pombe*)
 C; Species: *Schizosaccharomyces pombe*
 C; Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
 C; Accession: T39608
 R; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R. Submitted to the EMBL Data Library, May 1998
 A; Reference number: Z21866
 A; Accession: T39608
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Residues: 1-827 <LYN>
 A; Cross-references: EMBL:AL023554; PIDN:CAA19035.1; GSPDB:GN00067; SPDB:SPBC16G5.16
 A; Experimental source: strain 972h-; cosmid C16G5
 C; Genetics:
 A; Gene: SPDB:SPBC16G5.16
 A; Map position: 2
 A; Introns: 30/3; 41/1

C;Superfamily: GAL4 zinc binuclear cluster homology <GL4>
 F;11-47/Domain: GAL4 zinc binuclear cluster homology <GL4>

Query Match Score 36; DB 2; Length 827;
 Best Local Similarity 83.3%; Pred. No. 82;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 7
 Db 221 HIHHPA 226

RESULT 12

A82736 hypothetical protein XP1010 [imported] - *Xylella fastidiosa* (strain 9a5c)
 C;Species: *Xylella fastidiosa*
 C;Accession: A82736 #Sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 R;anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A;Reference number: A82515; MUID:20365717; PMID:10910347
 A;Note: for a complete list of authors see reference number A59328 below
 A;Accession: A82736
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-169 <SIM>
 A;Cross-references: GB:AE003938; GB:AE003849; NID:9105935; PIDN:AAF83820.1; GSPDB:GN0001

A;Experimental source: strain 9a5c
 R;Simpson, A.J.G.; Reinada, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carrer, E
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froeh
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A;Authors: Martins, E.M.F.; Matsuoka, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
 'F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, R.C.; Palmeira, R.V.; Sawasak
 A;Authors: Rodrigues, V.; Rosa, A.J.; de Rosa Oliveira, M.C.; de Oliveira, R.G.; Santelli, R.V.;
 A;Authors: da Silva, A.C.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
 A;Reference number: A59328
 A;Contents: annotation
 C;Genetics:

A;Gene: XF1010

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 A;Contents: annotation
 C;Genetics:

A;Gene: XF1010

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

Query Match Score 35; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 158 HIHHP 162

C;Genetics:
 A;Gene: pqqC; PA1987

Query Match Score 35; DB 2; Length 250;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 23 HIHHP 27

RESULT 14

S20455 pqqC protein - *Klebsiella pneumoniae*
 C;Species: *Klebsiella pneumoniae*
 C;Accession: S20455; S21840
 R;Meuleenberg, J.J.M.; Sellink, E.; Riegman, N.H.; Postma, P.W.
 Mol. Gen. Genet. 232, 284-294, 1992
 A;Title: Nucleotide sequence and structure of the *Klebsiella pneumoniae* pqq operon.
 A;Reference number: S20452; MUID:92212293; PMID:1313537
 A;Accession: S20455
 A;Molecule type: DNA
 A;Residues: 1-251 <MEU>
 A;Cross-references: EMBL:X58778; NID:943903; PIDN:CAA41581.1; PID:943907

C;Genetics:
 A;Gene: pqqC

Query Match Score 35; DB 2; Length 251;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 24 HIHHP 28

RESULT 15

A56051 myocyte nuclear factor - mouse
 C;Species: *Mus musculus* (house mouse)
 C;Accession: A56051
 R;Bassel-Duby, R.; Hernandez, M.D.; Yang, Q.; Seldin, J.M.; Rochelle, J.M.; Williams, R
 Mol. Cell. Biol. 14, 4596-4605, 1994
 A;Title: Myocyte nuclear factor, a novel winged-helix transcription factor under both d
 A;Reference number: A56051; MUID:94277065; PMID:8007964
 A;Accession: A56051
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-617 <BAS>
 A;Cross-references: GB:126507; NID:9508527; PIDN:AAA37529.1; PID:9508528

C;Genetics:
 A;Gene: MNF

C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
 C;Keywords: DNA binding; transcription factor
 F;289-380/Domain: fork head DNA-binding domain homology <FHD>

Query Match Score 35; DB 2; Length 617;
 Best Local Similarity 71.4%; Pred. No. 89;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHP 7
 Db 583 THHHP 589

RESULT 13

C83397 pyrroloquinoline quinone biosynthesis protein C PA1987 [imported] - *Pseudomonas aeruginosa*
 C;Species: *Pseudomonas aeruginosa*
 C;Accession: C83397 #Sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bi
 adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: C83397
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-250 <STO>
 A;Cross-references: GB:AE004625; PIDN:AA9947983; NID:99947983; PIDN:AG05375.1; GSPDB:GN001
 A;Experimental source: strain PA01

Search completed: February 11, 2004, 17:11:54
 Job time : 11.3333 secs

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL -!- FUNCTION: May play a role in transcriptional regulation
 CC interacting with ISWI. May serve a specific role in maintaining or
 CC altering the chromatin structure of the rDNA locus (By
 CC similarity).
 CC -!- SUBUNIT: Together with ISWI/SNF2h, it forms a complex termed NORC
 CC (nucleolar remodeling complex).
 CC -!- SUBCELLULAR LOCATION: Nuclear. Co-localizes with the basal RNA
 CC Polymerase I transcription factor UBF in the nucleolus.
 CC -!- TISSUE SPECIFICITY: Expressed at moderate levels in most tissues
 CC analyzed, including heart, brain, placenta, lung, skeletal muscle,
 CC kidney and pancreas.
 CC -!- SIMILARITY: BELONGS TO THE WAL FAMILY.
 CC -!- SIMILARITY: Contains 4 A-T hook DNA-binding repeats.
 CC -!- SIMILARITY: Contains 1 methyl-binding (MBD) domain.
 CC -!- SIMILARITY: Contains 1 bromodomain.
 CC -!- SIMILARITY: Contains 1 DDT domain.
 CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.
 CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -!- DR EMBL; AB032254; BAA89211.1; -.
 DR EMBL; AF000422; AAB60864.1; -.
 DR EMBL; AB002312; BAA20773.1; -.
 DR EMBL; BC008965; AAH08965.1; -.
 DR HSSP; Q92831; 1B91.
 DR Genew; HGNC:962; BAZ2A.
 DR MIM; 605682; -.
 DR GO; GO:0005731; C:nucleolus organizer complex; NAS.
 DR GO; GO:0003677; F:DNA binding activity; NAS.
 DR GO; GO:0030528; F:transcription regulator activity; NAS.
 DR GO; GO:0006338; P:chromatin modeling; NAS.
 DR InterPro; IPR000637; AT_hook.
 DR InterPro; IPR001487; Bromodomain.
 DR InterPro; IPR004022; DDT_dom.
 DR InterPro; IPR001739; Methyl-CpG_bind.
 DR InterPro; IPR001965; Znf_PHD.
 DR Pfam; PF02178; AT_hook; 4.
 DR Pfam; PF00439; bromodomain; 1.
 DR Pfam; PF02791; DDT; 1.
 DR Pfam; PF01429; MBD; 1.
 DR Pfam; PF00628; PHD; 1.
 DR PRINTS; PRO0503; BROMODOMAIN.
 DR SMART; SM00384; AT_hook; 4.
 DR SMART; SM00297; BROMO; 1.
 DR SMART; SM00571; DDT; 1.
 DR SMART; SM00391; MBD; 1.
 DR SMART; SM00249; PHD; 1.
 DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
 DR PROSITE; PS50014; BROMODOMAIN_2; 1.
 DR PROSITE; PS01359; ZF_PHD_1; FALSE_NEG.
 DR PROSITE; PS50016; ZF_PHD_2; 1.
 KW Transcription regulation; Bromodomain; Zinc-finger; Coiled coil;
 KW Nuclear protein; Repeat; DNA-binding.
 FT DOMAIN 525 577 MBD.
 FT DNA_BIND 622 634 A.T HOOK 1.
 FT DNA_BIND 643 655 A.T HOOK 2.
 FT DOMAIN 821 883 DDT.
 FT DNA_BIND 1159 1171 A.T HOOK 3.
 FT DNA_BIND 1377 1389 A.T HOOK 4.
 FT ZN_FING 1649 1699 PHD_TYPE.
 FT DOMAIN 1783 1853 BROMODOMAIN.

RA LYS-RICH.
 RA COILED COIL (POTENTIAL).
 RA GLU-RICH.
 RA PRO-RICH.
 RT POLY-ARG.
 RL L -> V (IN REF. 2).
 RL L -> Q (IN REF. 2 AND 3).
 RL H -> Q (IN REF. 2 AND 3).
 RL SKAEEKEKGKTKQ -> KIKKKKKKKKK (IN REF.
 RL 2).
 FT DOMAIN 633 772
 FT DOMAIN 666 765
 FT DOMAIN 1185 1250
 FT DOMAIN 1263 1384
 FT DOMAIN 1732 1735
 FT CONFLICT 574 574
 FT CONFLICT 700 700
 FT CONFLICT 720 720
 FT CONFLICT 727 738
 FT CONFLICT 785 785
 FT CONFLICT 951 951
 FT CONFLICT 1005 1006
 FT CONFLICT 1035 1037
 FT CONFLICT 1163 1163
 FT CONFLICT 1166 1166
 FT CONFLICT 1172 1172
 FT CONFLICT 1178 1178
 FT CONFLICT 1202 1202
 FT CONFLICT 1292 1292
 FT CONFLICT 1295 1295
 FT CONFLICT 1313 1313
 FT CONFLICT 1407 1410
 FT CONFLICT 1416 1416
 FT CONFLICT 1541 1541
 FT CONFLICT 1571 1571
 FT CONFLICT 1616 1616
 FT CONFLICT 1622 1622
 FT CONFLICT 1629 1629
 FT CONFLICT 1636 1636
 FT CONFLICT 1739 1739
 FT CONFLICT 1754 1754
 SQ SEQUENCE 1878 AA; 208480 MW; 400970CA6B234317 CRC64;
 SQ Query Match 86.4%; Score 38; DB 1; Length 1878;
 SQ Best Local Similarity 71.4%; Pred. No. 69;
 SQ Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
 BRN3_CHICK STANDARD; PRT; 341 AA.
 ID BRN3_CHICK
 AC Q91998;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Brain-specific homeobox/POU domain protein 3 (Brn-3).
 GN BRN3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TAXID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Dorsal root ganglion;
 RC MEDLINE=97349928; PubMed=9205808;
 RX Lindenberg J.K., Klint P., Williams R., Ebendal T.;
 RA "Identification of a chicken homologue in the Brn-3 subfamily of POU-
 RT transcription factors.";
 RL Brain Res. Dev. Brain Res. 100:169-182 (1997).
 CC -!- FUNCTION: MAY PLAY A ROLE IN SPECIFYING TERMINALLY DIFFERENTIATED
 CC NEURONAL PHENOTYPES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
 CC -!- CLASS-4 SUBFAMILY.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; X91998; CAA63049.1; -.
 DR EMBL; X91997; CAA63048.1; -.
 DR HSSP; P14859; LOC7.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000327; POU_domain.
 DR InterPro; IPR007103; POU_homeo.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00157; pou; 1.
 DR PRINTS; PRO0028; POU_DOMAIN.
 DR ProdDom; PD0000010; Homeobox; 1.
 DR ProdDom; PD0000583; POU_domain; 1.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00352; POU; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00035; POU_1; 1.
 DR PROSITE; PS00465; POU_2; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Transcription regulation; Nuclear protein.
 FT DOMAIN 55 65 POU-IV BOX.
 FT DOMAIN 185 259 POU.
 FT DNA_BIND 277 336 HOMEobox.
 SQ SEQUENCE 341 AA; 37773 MW; 6A4C741C88C07D5F CRC64;
 Query Match 84.1%; Score 37; DB 1; Length 341;
 Best Local Similarity 83.3%; Pred. No. 17;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HTHHPS 7
 Db 95 HTHHPS 100

RESULT 3
 ADD2_CAEEL STANDARD; PRT; 611 AA.
 AC Q20952;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adducin-related protein 2.
 GN ADD-2 OR F57F5.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderrinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN SEQUENCE FROM N.A.
 RA Moorthy S.; Bennett V.;
 RT "Molecular and functional analysis of the spectrin based membrane
 skeleton in *Caenorhabditis elegans*."
 RI Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
 RN SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Harris B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: MEMBRANE-CYTOSKELETON-ASSOCIATED PROTEIN THAT PROMOTES
 CC THE ASSEMBLY OF THE SPECTRIN-ACTIN NETWORK (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE ALDOLASE CLASS II FAMILY. ADDUCIN
 CC SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; AFI66171; AAD49860.1; -.
 DR EMBL; Z75953; CAB00100.1; -.
 DR PIR; T22855; T22855.
 DR WormPeP; F57F5.4; CE06002.
 DR InterPro; IPR001303; Aldolase_II_N.
 DR Pfam; PF00596; Aldolase_II; 1.
 KW Cytoskeleton; Membrane.
 SQ SEQUENCE 611 AA; 68561 MW; 9815707C45CD70D8 CRC64;
 Query Match 84.1%; Score 37; DB 1; Length 611;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HTHHPS 7
 Db 451 HTHHPS 456

RESULT 4
 HMN2_DROME STANDARD; PRT; 722 AA.
 ID HMN2_DROME
 AC P22808; Q24589;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Homeobox protein vnd (Ventral nervous system defective protein)
 DE (Homeobox protein NK-2).
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Arthropoda; Hexapoda; Insecta; Pterygota; Muscomorpha;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Drosophiliidae; Drosophila.
 OC Ephydriidae; Drosophiliidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95354667; PubMed=7628450;
 RA Jimenez F.; Martin-Morris L.E.; Velasco L.; Chu H.; Sierra J.;
 RA Rosen D.R.; White K.;
 RT "vnd, a gene required for early neurogenesis of Drosophila, encodes a
 homeodomain protein."
 RT homeodomain protein.";
 RL RN [1]
 RP SEQUENCE OF 476-633 FROM N.A.
 RX MEDLINE=90046666; PubMed=2573058;
 RA Kim Y.; Nirenberg M.;
 RT "Drosophila NK-homeobox genes."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7716-7720 (1989).
 RN SEQUENCE BY NMR OF 537-612.
 RP STRUCTURE BY NMR OF 537-612.
 RX MEDLINE=95371127; PubMed=7643404;
 RA Tsao D.H.H.; Gruschus J.M.; Wang L.-H.; Nirenberg M.; Ferretti J.A.;
 RT "The three-dimensional solution structure of the NK-2 homeodomain
 from *Drosophila*."
 RL J. Mol. Biol. 251:297-307 (1995).
 RN STRUCTURE BY NMR OF 543-605.
 RP STRUCTURE BY NMR OF 543-605.
 RX MEDLINE=97299770; PubMed=9154919;
 RA Gruschus J.M.; Tsao D.H.H.; Wang L.-H.; Nirenberg M.; Ferretti J.A.;
 RT "Interactions of the vnd/NK-2 homeodomain with DNA by nuclear magnetic
 resonance spectroscopy: basis of binding specificity."
 RL Biochemistry 36:5372-5380 (1997).
 CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN THE
 CC REGULATION OF THE PRONEURAL AS-C GENES AND THE NEUROGENIC GENES OF
 CC THE ENHANCER OF SPLIT COMPLEX. COULD SPECIFICALLY ACTIVATE
 CC PRONEURAL GENES IN THE VENTRAL-MOST NEUROECTODERM.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE CNS AND MIDGUT.
 CC -!- SIMILARITY: BELONGS TO THE NK-2 HOMEBOX FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X87141; CAA60619.1; -.
 DR PIR; M27290; AAA28617.1; -.
 DR PIR; S57246; S57246.
 DR PDB; 1VND; 0B-NOV-96.
 DR PDB; 1NK2; 23-FEB-99.
 DR PDB; 1NK3; 23-FEB-99.
 DR PDB; 1QRY; 06-JUL-99.
 DR TRANSFAC; T04258; -.
 DR FlyBase; FBgn0003986; vnd.
 DR GO; GO:0007400; P:neuroblast cell fate determination; IMP.
 DR GO; GO:0007419; P:ventral cord development; NAS.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRODOM; PD0000010; Homeobox; 1.
 DR SMART; SM000389; HOX; 1.
 DR PROSITE; PS00027; HOMEobox¹; 1.
 DR PROSITE; PS50071; HOMEobox²; 1.
 KW transcription regulation; Homeobox; DNA-binding; 3D-structure;
 KW Developmental protein; Nuclear protein; Neurogenesis.
 FT DOMAIN 108 111 POLY-SER.
 FT DOMAIN 215 223 POLY-ALA.
 FT DOMAIN 336 343 POLY-ALA.
 FT DOMAIN 351 359 POLY-ALA.
 FT DOMAIN 482 486 POLY-ASN.
 FT DOMAIN 489 492 POLY-ASN.
 FT DNA_BIND 544 603 HOMEobox.
 FT DOMAIN 688 693 POLY-ALA.
 FT VARIANT 578 578 A -> T (IN VND29).
 FT CONFLICT 632 633 RR -> VG (IN REF. 2).
 FT TURN 553 554
 FT HELIX 555 565
 FT HELIX 571 581
 FT TURN 582 582
 FT HELIX 585 594
 SQ SEQUENCE 722 AA; 76468 MW; D036AE4D890014DA CRC64;

Query Match 84.1%; Score 37; DB 1; Length 722;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 1 THIHHP 7
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 1 THIHHP 7
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 81.8%; Score 36; DB 1; Length 376;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN NCBI_TaxID=33703;

Query Match 81.8%; Score 36; DB 1; Length 376;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN NCBI_TaxID=33703;

CC -!- SIMILARITY: BELONGS TO THE RLU FAMILY OF PSEUDOURIDINE SYNTHASES.
 CC -!- STRONG, TO E.COLI YQCB.
 CC -!- SIMILARITY: IN THE N-TERMINAL, STRONG, TO E.COLI YQCB.
 CC -!- TERMINAL, STRONG, TO E.COLI YQCB.
 CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration - between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -!- EMBL; X79474; CAA55982.1; -.
 DR PIR; S45107; S45107.
 DR InterPro; IPR006145; PseudoU synth.
 DR InterPro; IPR006224; PSI_RLU.
 DR Pfam; PF04287; DUFE46; 1.
 DR Pfam; PF00849; PseudoU synth_2; 1.
 DR PRODOM; PD001819; PSI_RLU; 1.
 DR PROSITE; PS01129; PSI_RLU; 1.
 SQ SEQUENCE 376 AA; 43605 MW; R35992CAAAD22E30 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 376;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN NCBI_TaxID=33703;

RESULT 6
 VPAP_PRVKA
 ID VPAP_PRVKA STANDARD; PRT; 384 AA.
 AC P36702;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE DNA polymerase processivity factor (Polymerase accessory protein)
 DE (PAP) (UL42 homolog).
 DE Pseudorabies virus (strain Kaplan) (PRV).
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN NCBI_TaxID=33703;

SEQUENCE FROM N.A.
 RX MEDLINE=9522277; PubMed=7707503;
 RA Berthomme H., Monahan S.J., Parris D.S., Jacquemont B.,
 RA Epstein A.L.;
 RT RT subunits of the pseudorabies virus DNA polymerase holoenzyme:
 RT evidence for specificity of interaction.";
 RL J. Virol. 69:2811-2818 (1995).
 CC -!- FUNCTION: ACCESSORY SUBUNIT OF THE DNA POLYMERASE THAT ACTS TO
 CC INCREASE THE PROCESSIVITY OF POLYMERIZATION (BY SIMILARITY).
 CC -!- SIMILARITY: TO OTHER HERPESVIRUSES POLYMERASE ACCESSORY PROTEIN.
 CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration - between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -!- EMBL; M94355; AAA74384.1; -.
 DR InterPro; IPR003202; UL42.
 DR Pfam; PF02282; UL42; 2.
 KW DNA-binding; DNA replication.
 SQ SEQUENCE 384 AA; 40305 MW; RFAC3EF0C1984936 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 384;

Best Local Similarity 71.4%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1	THIHPS	7
16	THAHHPA	22

RESULT 7

3CAA_CHICK STANDARD; PRT; 637 AA.

3C Q92075; P70095; Q98941; 01-NOV-1997 (Rel. 35, Created) 01-NOV-1997 (Rel. 35, Last sequence update) 28-FEB-2003 (Rel. 41, Last annotation update)

Amiloride-sensitive sodium channel alpha-subunit (Epithelial Na⁺ channel alpha subunit) (Alpha ENaC) (Nonvoltage-gated sodium channel 1 alpha subunit) (SCNEA) (Alpha NaCh).

SCNN1A OR ENAC.

Gallus gallus (Chicken).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

NCBI_TaxID=9031;

[1] SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
 STRAIN=Isa brown; TISSUE=Cochlea;
 MEDLINE=97157073; PubMed=9003454;
 Killick R., Richardson G.; "Isolation of chicken alpha ENaC splice variants from a cochlear cDNA library." Biochim. Biophys. Acta 1350:33-37 (1997).

[2] SEQUENCE FROM N.A.
 TISSUE=Intestine;
 Goldstein O., Asher C., Garty H.; Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: SODIUM PERMEABLE NON-VOLTAGE-SENSITIVE ION CHANNEL INHIBITED BY THE DIURETIC AMILORIDE. MEDIATE THE ELECTRODIFFUSION OF THE LUMINAL SODIUM (AND WATER, WHICH FOLLOWS OSMOTICALLY) THROUGH THE APICAL MEMBRANE OF EPITHELIAL CELLS. CONTROLS THE REABSORPTION OF SODIUM IN KIDNEY, COLON, LUNG AND SWEAT GLANDS. ALSO PLAYS A ROLE IN TASTE PERCEPTION (BY SIMILARITY).
 -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA, ONE BETA AND ONE GAMMA SUBUNIT. A DELTA SUBUNIT CAN REPLACE THE ALPHA SUBUNIT (BY SIMILARITY).
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=Long; Name=Short;

IsoID=Q92075-1; Sequence=Displayed;

[3] ISOFORM=Q92075-2; Sequence=VSP 006196, VSP 006197;

-!- TISSUE SPECIFICITY: THE LONG ISOFORM HAS BEEN FOUND IN COCHLEA, COLON, AND CARTILAGE. THE SHORT ISOFORM IS ONLY FOUND IN COCHLEA.
 -!- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNEL FAMILY.
 -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-15 IS THE INITIATOR.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

EMBL; U62902; AAB50550.1; -
 EMBL; U62903; AAB50551.1; -
 EMBL; U62904; AAB50552.1; -
 EMBL; US8475; AAB04954.1; ALIT_INIT.
 InterPro; IPR004724; EnaC.
 InterPro; IPR001873; Natchannel ASC.

Pfam; PF00858; ASC; 1.	DR	PRINTS; PRO1078; AMINACHANNEL
TIGRFAMS; TIGR00859; ENaC; 1.	DR	
PROSITE; PS01206; ASC; 1.	DR	
Ionic channel; Transmembrane;	KW	
Alternative splicing;	KW	
DOMAIN 1 80	FT	
TRANSMEM 81 97	FT	
DOMAIN 98 554	FT	
TRANSMEM 555 571	FT	
DOMAIN 572 637	FT	
CARBOHYD 157 157	FT	
CARBOHYD 280 280	FT	
CARBOHYD 298 298	FT	
CARBOHYD 499 499	FT	
VARSPLIC 402 434	FT	
VARSPLIC 435 637	FT	
CONFLICT 49 50	FT	
CONFLICT 578 578	FT	
SEQUENCE 637 AA; 73236 MW;	SQ	
Query Match 81.8%;		
Best Local Similarity 83.3%;		
Matches 5; Conservative 0		
Qy 1 THIFHP 6		
Db 282 THFHP 287		

Query Match 79.5%; Score 35; DB 1; Length 251;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 HIHHP 6
 Db 24 HIHHP 28

RESULT 9
 FXK1_MOUSE STANDARD; PRT; 617 AA.

ID FXK1_MOUSE; STANDARD; PRT; 617 AA.

AC P42128; O35939;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Forkhead box protein X1 (Myocyte nuclear factor) (MNF).
 GN FOXK1 OR MNF.
 OS Mus musculus (Mouse).
 OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]

SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=94277065; PubMed=8007964;
 RA Bassel-Duby R.; Hernandez M.D.; Yang Q.; Rochelle J.M.,
 RA Seldin M.F.; Williams R.S.;
 RT "Myocyte nuclear factor, a novel winged-helix transcription factor
 under both developmental and neural regulation in striated
 myocytes.";
 RT Mol. Cell. Biol. 14:4596-4605 (1994).
 RL [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97415602; PubMed=9271401;
 RA Yang Q.; Bassel-Duby R.; Williams R.S.;
 RT "Transient expression of a winged-helix protein, MNF-beta, during
 RT myogenesis.";
 RL Mol. Cell. Biol. 17:5236-5243 (1997).
 CC -- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO THE UPSTREAM
 CC ENHANCER REGION (CCAC BOX) OF MYOGLOBIN GENE. HAS A ROLE IN
 CC MYOGENIC DIFFERENTIATION AND IN REMODELING PROCESSES OF ADULT
 CC MUSCLES THAT OCCUR IN RESPONSE TO PHYSIOLOGICAL STIMULI.
 CC -- SUBCELLULAR LOCATION: Nuclear.
 CC -- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoID=P42128-1; Sequence=Displayed;
 CC Name=2; Synonyms=Beta;
 CC IsoID=P42128-2; Sequence=VSP 001545; VSP 001546;
 CC -- TISSUE SPECIFICITY: EXPRESSED IN TISSUES AND CELLS IN WHICH THE
 CC MYOGLOBIN GENE IS TRANSCRIPTIONALLY ACTIVE (CARDIAC AND SKELETAL
 CC MYOCYTES, BRAIN, KIDNEY. .).
 CC -- PTM: PHOSPHORYLATED.
 CC -- SIMILARITY: Contains 1 fork-head domain.
 CC -- SIMILARITY: Contains 1 FHA domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; L26507; AAA37529.1; -
 DR EMBL; U95016; AAB69641.1; -
 DR PIR; A56051; A56051.
 DR HSSP; Q63245; 2HFP.
 DR TRANSFAC; T04216; -
 DR MGD; MGI:1347488; Foxk1.
 DR InterPro; IPR000253; Fha.
 DR InterPro; IPR001766; Fha.
 DR Pfam; PF00498; Fha; 1.

DR SMART; SMO0339; Fha; 1.
 DR PROSITE; PS000240; Fha; 1.
 DR PROSITE; PS000657; FOK HEAD_1; 1.
 DR PROSITE; PS000658; FOK HEAD_2; 1.
 DR PROSITE; PS50039; FOK HEAD_3; 1.
 DR PROSITE; PS50006; FHA DOMAIN; 1.
 KW DNA-binding; Transcription regulation; Activator; Nuclear protein;
 KW Phosphorylation; Alternative splicing; Activator; Differentiation.
 FT DOMAIN 27 33 POLY-ALA.
 FT DOMAIN 107 159 FHA.
 FT DNA BIND 288 379 FORK-HEAD.
 FT VARSPLIC 400 409 RSPASPTHPG -> SAPASHTSHA (in isoform 2).
 FT VARSPLIC 410 617 /FTId=VSP 001545.
 FT VARSPLIC 410 617 Missing (in isoform 2).
 SQ SEQUENCE 617 AA; 65839 MW; A1083B28C709FC4A CRC64;
 SQ SEQUENCE 617 AA; 79.5%; Score 35; DB 1; Length 617;
 Best Local Similarity 71.4%; Pred. No. 69;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIHP 7
 Db 583 THSHRPT 589

RESULT 10
 UL47_HSV11 STANDARD; PRT; 693 AA.

ID UL47_HSV11 STANDARD; PRT; 693 AA.

AC P10231;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE Virion protein UL47 (82/81 kDa tegument protein) (VMW82/81) (VP13/14).
 GN UL47.
 OS Herpes simplex virus (type 1 / strain 17).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=68274327; PubMed=2839594;
 RA McGeoch D.J.; Dalrymple M.A.; Davison A.J.; Dolan A.; Frame M.C.,
 RA McNab D.; Perry L.J.; Scott J.E.; Taylor P.;
 RT "The complete DNA sequence of the long unique region in the genome of
 RT herpes simplex virus type 1."
 RL Gen. Virol. 69:1531-1574 (1988).
 RN [2]
 RP IDENTIFICATION OF PROTEIN.
 RX MEDLINE=91108391; PubMed=2177077;
 RA McLean G.; Rixon F.; Langeland N.; Haarr L.; Marsden H.;
 RT "Identification and characterization of the virion protein products
 RT of herpes simplex virus type 1 gene UL47."
 RL Gen. Virol. 71:2953-2960 (1990).
 CC -- FUNCTION: MODULATOR OF ALPHA-TIF (VMW65 PHOSPHOPROTEIN) TRANS-
 CC -- ACTIVATION. POSSIBLY UL47 MAY HAVE KINASE ACTIVITY.
 CC -- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
 CC -- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
 CC -- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
 CC EHV-1 13, AND VZV 11.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC DR EMBL; L26507; AAA37529.1; -
 CC DR EMBL; U95016; AAB69641.1; -
 CC DR PIR; A56051; A56051.
 CC DR HSSP; Q63245; 2HFP.
 CC DR TRANSFAC; T04216; -
 CC DR MGD; MGI:1347488; Foxk1.
 CC DR InterPro; IPR000253; Fha.
 CC DR InterPro; IPR001766; Fha.
 CC DR Pfam; PF00498; Fha; 1.

-!- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
 CC DISTRIBUTION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: J05166; AAA40799.1; -
 DR PIR: A34911; A34911.
 DR HSSP; P02730; IBTO.
 DR InterPro; IPR001717; Anion exchange.
 DR InterPro; IPR003020; HCO3_cotransp.
 DR Pfam; PF00955; HCO3_cotransp; 1.
 DR PRINTS; PRO1231; HCO3TRANSPORT.
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS002119; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS002220; ANION_EXCHANGER_2; 1.
 KW Transmembrane; Glycoprotein; Transport; Antiport; Ion transport;
 KW Anion exchange; Lipoprotein; Palmitate.
 FT DOMAIN 1 704 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 705 1234 MEMBRANE (ANION EXCHANGE).
 FT TRANSMEM 705 728 POTENTIAL.
 FT TRANSMEM 734 771 POTENTIAL.
 FT TRANSMEM 791 813 POTENTIAL.
 FT TRANSMEM 823 844 EXOPLASMIC LOOP (POTENTIAL).
 FT DCMAIN 844 893 POTENTIAL.
 FT TRANSMEM 894 911 POTENTIAL.
 FT DOMAIN 912 926 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 927 947 POTENTIAL.
 FT TRANSMEM 981 1003 POTENTIAL.
 FT TRANSMEM 1029 1050 POTENTIAL.
 FT TRANSMEM 1084 1129 POTENTIAL.
 FT TRANSMEM 1156 1192 POTENTIAL.
 FT DOMAIN 5 317 PRO-RICH.
 FT CARBOHYD 74 88 HIS-RICH.
 FT CARBOHYD 856 856 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 878 878 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT LIPID 1166 1166 PALMITATE (BY SIMILARITY).
 FT CONFLICT 206 206 G -> A (IN REF. 2).
 FT CONFLICT 925 926 RR -> PG (IN REF. 2).
 FT CONFLICT 1018 1018 M -> I (IN REF. 2).
 FT CONFLICT 1156 1156 M -> I (IN REF. 2).
 SQ SEQUENCE 1234 AA; 1366355 MW; FAB4ED12BB916216 CRC64;
 Query Match 79.5%; Score 35; DB 1; Length 1234;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 80 HIHHP 84

RESULT 13
 ID B3A2_MOUSE STANDARD: PRT: 1237 AA.
 AC P13808; Q9ES09; Q9ES10; Q9ES11; Q9ES12; Q9ES13;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Anion exchange protein 2 (Non-erythroid band 3-like protein) (B3RP).
 GN SLC4A2 OR AE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;

[1] RN SEQUENCE FROM N.A. (ISOFORM A).
 RP MEDLINE=89034212; PubMed=3182834;
 RX Alper S.L., Kopito R.R., Libresco S.M., Lodish H.F.;
 RA "Cloning and characterization of a murine band 3-related cDNA from
 RT kidney and from a lymphoid cell line.";
 RL J. Biol. Chem. 263:17092-17099 (1988).
 RN [2] RN SEQUENCE FROM N.A. ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RP MEDLINE=20462926; PubMed=11006093;
 RX Lecanda J., Urtasun R., Medina J.F.;
 RA "Molecular cloning and genomic organization of the mouse AE2 anion
 RT exchanger gene.";
 RL Biochem. Biophys. Res. Commun. 276:117-124 (2000).
 CC -!- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
 CC DISTRIBUTION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS;
 CC Event=Alternative splicing; Named isoforms=S;
 CC Name=A;
 CC IsoID=P13808-1; Sequence=Displayed;
 CC Name=B1;
 CC IsoID=P13808-2; Sequence=VSP_000458;
 CC Name=B2;
 CC IsoID=P13808-3; Sequence=VSP_000457;
 CC Name=C1;
 CC IsoID=P13808-4; Sequence=VSP_000460;
 CC Name=C2;
 CC IsoID=P13808-5; Sequence=VSP_000459; VSP_000461;
 CC -!- TISSUE SPECIFICITY: Isoform a is widely expressed at similar
 CC levels in all tissues examined. Isoforms b1 and b2 are
 CC predominantly expressed in stomach although they are also detected
 CC at lower levels in other tissues. Isoform c1 is stomach-specific.
 CC Isoform c2 is expressed at slightly higher levels in lung and
 CC stomach than in other tissues.
 CC -!- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR J04036; AAA65505.1; -
 DR EMBL; AF255774; AAG23154.1; -
 DR EMBL; AF255774; AAG23155.1; -
 DR EMBL; AF255774; AAG23156.1; -
 DR EMBL; AF255774; AAG23158.1; -
 DR EMBL; AF255774; AAG23157.1; -
 DR PIR; A31789; A31789.
 DR HSSP; P02730; IBTO.
 DR MGI; 109351; S1c4a2.
 DR InterPro; IPR001717; Anion exchange.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR InterPro; IPR001717; Anion exchange.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane (ANION EXCHANGE).
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 KW Transmembrane; Glycoprotein; Transport; Antipart; Ion transport;
 KW Transmembrane

CC EMBL; AF121253; AAD19700.1; .
 DR HSSP; P02730; 1BTQ.
 DR InterPro; IPR001717; Anion_exchange.
 DR InterPro; IPR003020; HCO3_cotranspt.
 DR Pfam; PF00955; HCO3_cotransport; 1.
 DR PRINTS; PRO1231; HC03_TRANSPORT; 1.
 DR TIGRFAMS; TIGR00834; ae; 1.
 DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
 DR Transmembrane; Glycoprotein; Transport; Antiport; Ion transport;
 KW Anion exchange; Lipoprotein; Palmitate.
 KW Anion exchange; Lipoprotein; Palmitate.
 DOMAIN 1 704 CYTOPLASMIC (POTENTIAL).
 DOMAIN 705 1238 MEMBRANE (ANION EXCHANGE).
 FT TRANSMEM 705 728 POTENTIAL.
 FT TRANSMEM 734 771 POTENTIAL.
 FT TRANSMEM 791 813 POTENTIAL.
 FT TRANSMEM 823 844 POTENTIAL.
 FT DOMAIN 845 897 EXOPLASMIC LOOP (POTENTIAL).
 FT TRANSMEM 898 915 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 916 930 POTENTIAL.
 FT TRANSMEM 931 951 POTENTIAL.
 FT TRANSMEM 985 1007 POTENTIAL.
 FT TRANSMEM 1033 1054 POTENTIAL.
 FT TRANSMEM 1088 1133 POTENTIAL.
 FT TRANSMEM 1160 1196 POTENTIAL.
 FT DOMAIN 5 316 PRO-RICH.
 FT DOMAIN 74 88 HIS-RICH.
 FT CARBOHYD 856 856 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 865 865 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 879 879 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT LIPID 1170 1170 PALMITATE (BY SIMILARITY).
 SQ SEQUENCE 1238 AA; 137358 MW; FA1739862ED5ADBF CRC64;

Query Match Score 35; DB 1; Length 1238;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHP 6
 Db 80 HIHHP 84

Search completed: February 11, 2004, 17:04:16
 Job time : 7.16667 secs

Copyright GenCore version 5.1.6 (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

Run on: February 11, 2004, 16:52:34 ; Search time 24.58333 Seconds
(without alignments)
73.479 Million cell updates/sec

Writitle: US-09-901-187C-11
Userperfect score: 44
Sequence: 1 THIHHP\$ 7

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

SPTREMBL 23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp Rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriaph: *
17: sp_prochlorococcus: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	4.2	95.5	468	10	081027	081027 arabidops
2	4.0	90.9	1202	5	Q8IDT1	Q8IDT1 plasmodium
3	3.9	88.6	73	5	Q8SY32	Q8SY32 drosophila
4	3.9	88.6	312	10	Q9C7M4	Q9C7M4 arabidops
5	3.9	88.6	402	2	Q53371	Q53371 escherichia
6	3.9	88.6	402	2	Q8VSK3	Q8VSK3 shigella f
7	3.9	88.6	402	2	Q9K388	Q9K388 salmonella
8	3.9	88.6	402	2	Q8RQK7	Q8RQK7 bacillus c
9	3.9	88.6	402	2	Q56368	Q56368 escherichia
10	3.9	88.6	402	2	Q8GCG5	Q8GCG5 citrobacte
11	3.9	88.6	402	4	Q9H5N6	Q9H5N6 homo sapien
12	3.9	88.6	402	4	Q9BTF6	Q9BTF6 homo sapien
13	3.9	88.6	402	5	Q95TWO	Q95TWO drosophila
14	3.9	88.6	402	5	Q95T96	Q95T96 drosophila
15	3.9	88.6	402	12	Q8QTC0	Q8QTC0 white spo
16	3.9	88.6	402	12	Q9BAP4	Q9BAP4 bluetongu

ALIGNMENTS					
RESULT 1					
081027	ID	Q81027; Q9PFPE2; PRELIMINARY;	PRT;	468	AA.
AC	081027; Q9PFPE2;				
DT	01-NOV-1998 (TREMBLrel. 08, Created)				
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Putative hydroxymethylglutaryl-CoA lyase (Putative				
DE	hydroxymethylglutaryl-CoA lyase protein).				
GN	AT2G26800.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes; core eudicots; Arabidopsis.				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; Brassicaeae; Arabidopsis.				
OC	eurosid II; Brassicales; Brassicaceae;				
OX	NCBI_TaxID=3702;				
RN	[1]	SEQUENCE FROM N.A.			
RP	STRAIN=cv. Columbia;				
RC	Roundsley S.D., Ronning C.M., Lin X., Ketchum K.A., Crosby M.L.				
RA	Brandon R.C., Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R.				
RA	Adams M.D., Somerville C.R., Venter J.C.;				
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.				
RN	[2]	SEQUENCE FROM N.A.			
RP	STRAIN=cv. Columbia;				
RC	Lin X.;				
RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
RL	[3]	SEQUENCE FROM N.A.			
RN	SEQUENCE FROM N.A.				
RP	STRAIN=cv. Columbia;				
RC	Town C.D., Kaul S.;				
RA	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.				
RL	[4]	SEQUENCE FROM N.A.			
RN	SEQUENCE FROM N.A.				
RP	STRAIN=cv. Columbia;				
RA	Sakano H., Liu S.X., Pham P.K., Yamada K., Banh J., Etgui P.,				
RA	Toriumi M., Yu G., Shinozaki K., Davis R.W., Ecker J.R., Theo				
RT	"Full Length cDNA of gene F12C20_16/At2g26800 (GI:3426048)." ;				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
RN	[5]	SEQUENCE FROM N.A.			

SEQUENCE FROM N.A.
 RP Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 RA Carninici P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda A., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full Length cDNA of gene F12C20_16/At2g26800 (GI:3426048)." ;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005168; AAC32247.2; -.
 DR EMBL; AF327420; AAG42010.1; -.
 DR EMBL; AF349521; AAKL5568.1; -.
 DR InterPro; IPR000891; HMGL-like.
 DR Pfam; PF00682; HMGL-like; 1.
 KW Lyase.
 SQ SEQUENCE 468 AA; 50577 MW; C91C70DAD241AF74 CRC64;
 Query Match 95.5%; Score 42; DB 10; Length 468;
 Best Local Similarity 85.7%; Pred. No. 14;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 THIHHP 7
 Db 21 THIHHP 27
 RESULT 2
 Q8IDT1 ID Q8IDT1; PRELIMINARY; PRT; 1202 AA.
 AC DT 01-MAR-2003 (TREMBLrel. 23, Created)
 AC DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 AC DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE DE Hypothetical protein.
 GN GN PF13_0221.
 OS OS Plasmodium falciparum (isolate 3D7).
 OC OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21016719; PubMed=11130712;
 RX RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA RA White O., Alonso J., Araujo R., Bowman C.L., Brooks S.Y.,
 RA RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA RA Chung M.K., Conn L., Conway A.R., Creasy T.H., Dewar K.,
 RA RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.P., Hughes B., Huizar L.,
 RA RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA RA Kim C.J., Koo H.J., Kremmenetskaya I., Kurtz D.B., Kwan A., Lam B.,
 RA RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA RA Lin X., Liu S.X., Liu Z.A., Luors J.S., Maiti R., Marziali A.,
 RA RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA RA Utterback T., Van Aken S., Vaysberg M., Vysotskia V.S., Walker M.,
 RA RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820 (2000).
 DR DR EMBL; AC0068901; AAG50897.1; -.
 DR DR InterPro; IPR002559; Transposase_11_1.
 DR DR Pfam; PF01609; Transposase_11_1.
 SQ SEQUENCE 35554 MW; 3C06A8F768CE29EB3 CRC64;
 Query Match 88.6%; Score 39; DB 10; Length 312;
 Best Local Similarity 85.7%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 THIHHP 6
 Db 294 THIHHP 299
 RESULT 3
 Q8SY32 ID Q8SY32; PRELIMINARY; PRT; 73 AA.
 AC AC Q8SY32; Created)
 DT DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT DT 01-OCT-2002 (TREMBLrel. 21, Last annotation update)
 DE DE LP02734P.
 GN GN BCDNA; LP02734.
 OS OS Drosophila melanogaster (Fruit fly).
 OC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC OC Ephydrioidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]

RESULT 5

Q53371 PRELIMINARY; PRT; 402 AA.

ID Q53371; PRELIMINARY; PRT; 402 AA.

AC Q53371; 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Transposase (IS10 transposase).

GN R0085.

OS Escherichia coli, and

OG Salmonella typhi.

OC Plasmid pXT107, and Plasmid R27.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562, 601;

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES=E.coli; PLASMID=pXT107;

RX MEDLINE=94040791; PubMed=8224890;

RA Bogosian G., Bilyeu K., O'Neil J.P.;

RT "Genome rearrangements by residual IS10 elements in strains of

RT Escherichia coli K-12 which had undergone Tn10 mutagenesis and fusaric

RT acid selection.";

RT Gene 133:17-22 (1993).

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES=S.typhi; PLASMID=R27;

RX MEDLINE=20280091; PubMed=10773089;

RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,

RA Grotbeck E., Rose D.J., Taylor D.E.;

RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid

RT from *Salmonella typhi* that is temperature sensitive for transfer.";

RT Nucleic Acids Res. 28:2177-2186 (2000).

RN [3]

RP SEQUENCE FROM N.A.

RC SPECIES=S.typhi; PLASMID=R27;

RX MEDLINE=20247426; PubMed=10783303;

RA Lawley T.D., Burland V.D., Taylor D.E.;

RT "Analysis of the complete nucleotide sequence of the tetracycline-

RT resistance transposon Tn10.";

RL Plasmid 43:235-239 (2000).

DR EMBL; S67119; AAB28848.1; -

DR EMBL; AF250878; AAF69923.1; -

DR EMBL; AF223162; AAF68939.1; -

DR InterPro; IPR002559; Transposase_11.

DR Pfam; PF01609; Transposase_11; 1.

KW Plasmid.

SQ SEQUENCE 402 AA; 46041 MW; 2B4AB708C885667C CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;

Best Local Similarity 85.7%; Pred. No. 39;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 247 THCHHPS 253

RESULT 6

Q8VSK3 PRELIMINARY; PRT; 402 AA.

ID Q8VSK3; PRELIMINARY; PRT; 402 AA.

AC Q8VSK3; 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Transposase.

GN CP0084; *Shigella flexneri* 2a.

OS Plasmid pCP301.

OG Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; *Shigella*.

OX NCBI_TaxID=42897;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=301;

RA Jin Q., Zhang J.Y., Liu H., Yang J.J., Yang F., Zhang X.B., Wang J.H.,

RA Yang G.W., Wu H.T., Dong J., Sun L.L., Xue Y., Zhao A.L., Gao Y.S.,

RA Zhu J.P., Kan B., Chen S.X., Yao Z.J., He B.K., Chen R.S., Ma D.L.,

RA Yuan Z.H., Xu J.G., Wang Y., Shen Y., Lu W.C., Qiang B.Q., Wen Y.M.,

RA Hou Y.D.;

RT "Complete DNA sequence and analysis of the large virulence plasmid

RT PCP301 of *Shigella flexneri*.";

RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF386526; AAL72416.1;

DR InterPro; IPR002559; Transposase_11.

KW Plasmid.

SQ SEQUENCE 402 AA; 46046 MW; 1E9391EE324A12B6 CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;

Best Local Similarity 85.7%; Pred. No. 39;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPs 7

Db 247 THCHHPS 253

RESULT 7

Q9K388 PRELIMINARY; PRT; 402 AA.

ID Q9K388; PRELIMINARY;

AC Q9K388;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Putative IS10 transposase.

RN R0076.

OS *Salmonella typhi*.

OG Plasmid R27.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; *Salmonella*.

NCBI_TaxID=601;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20280091; PubMed=10773089;

RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,

RA Grotbeck E., Rose D.J., Taylor D.E.;

RA "The complete DNA sequence and analysis of R27, a large IncHI plasmid

RT from *Salmonella typhi* that is temperature sensitive for transfer.";

RT Nucleic Acids Res. 28:2177-2186 (2000).

RL R0076.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20247426; PubMed=10783303;

RA Lawley T.D., Burland V.D., Taylor D.E.;

RT "Analysis of the complete nucleotide sequence of the tetracycline-

RT resistance transposon Tn10.";

RL Plasmid 43:235-239 (2000).

DR EMBL; S67119; AAB28848.1; -

DR EMBL; AF250878; AAF69923.1; -

DR EMBL; AF223162; AAF68939.1; -

DR InterPro; IPR002559; Transposase_11.

DR Pfam; PF01609; Transposase_11; 1.

KW Plasmid.

SQ SEQUENCE 402 AA; 46059 MW; 25854FABD793130F CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;

Best Local Similarity 85.7%; Pred. No. 39;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPs 7

Db 247 THCHHPS 253

RESULT 8

Q8RQK7 PRELIMINARY; PRT; 402 AA.

ID Q8RQK7; PRELIMINARY;

AC Q8RQK7; DT 01-JUN-2002 (TREMBLrel. 21; Created) DT 01-JUN-2002 (TREMBLrel. 21; Last sequence update) DE Transposase. GN Bacillus cereus. OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus. OX NCBI_TAXID=1396; RN SEQUENCE FROM N.A. STRAIN=Tim-r01; RC SPECIES=S. cereus. RA Sampei G., Mizobuchi K.; RT "Organization and diversification of plasmid genomes: complete nucleotide sequence of the R100 genome."; RT Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases. [1] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Sampei G., Komano T., Sasaki T., Tachibana K., Furuya N., Saito Y., RA Suzuki T., Mizobuchi K.; RT "Organization and diversification of plasmid genomes: complete nucleotide sequence of R64 genome."; RT Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases. [2] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Komano T., Kubo A., Nisioka T.; RT "Shufflon: multi-inversion of four contiguous DNA segments of plasmid R64 creates seven different open reading frames."; RT Nucleic Acids Res. 15:1165-1172(1987). [3] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Kubo A., Kusukawa A., Komano T.; RT "Nucleotide sequence of the rci gene encoding shufflon-specific DNA recombinase in the IncII plasmid R64: homology to the site-specific recombinases of integrase family."; RT Mol. Gen. Genet. 213:30-35(1988). [4] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Komano T., Toyoshima A., Morita K., Nisioka T.; RT "Cloning and nucleotide sequence of the oriT region of the IncII plasmid R64."; RT J. Bacteriol. 170:4385-4387(1988). [5] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Furuya N., Nisioka T., Komano T.; RT "Nucleotide sequence and functions of the oriT operon in IncII plasmid R64."; RT J. Bacteriol. 173:2231-2237(1991). [6] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Kim S.R., Funayama N., Komano T.; RT "Determination of the nick site at oriT of IncII plasmid R64: global similarity of oriT structures of IncII and IncP plasmids."; RT J. Bacteriol. 175:5035-5042(1993). [7] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Furuya N., Komano T.; RT "Surface exclusion gene of IncII plasmid R64: nucleotide sequence and analysis of deletion mutants."; RT Plasmid 32:80-84(1994).

RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases. [5] RN SEQUENCE FROM N.A. RP SPECIES=S. coli; PLASMID=R100; RC SPECIES=S. coli; PLASMID=R100; RA Sampei G., Mizobuchi K.; RT "Organization and diversification of plasmid genomes: complete nucleotide sequence of the R100 genome."; RT Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases. [6] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Sampei G., Komano T., Sasaki T., Tachibana K., Furuya N., Saito Y., RA Suzuki T., Mizobuchi K.; RT "Organization and diversification of plasmid genomes: complete nucleotide sequence of R64 genome."; RT Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases. [7] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Komano T., Kubo A., Nisioka T.; RT "Shufflon: multi-inversion of four contiguous DNA segments of plasmid R64 creates seven different open reading frames."; RT Nucleic Acids Res. 15:1165-1172(1987). [8] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Kubo A., Kusukawa A., Komano T.; RT "Nucleotide sequence of the rci gene encoding shufflon-specific DNA recombinase in the IncII plasmid R64: homology to the site-specific recombinases of integrase family."; RT Mol. Gen. Genet. 213:30-35(1988). [9] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S typhimurium; PLASMID=R64; RA Komano T., Toyoshima A., Morita K., Nisioka T.; RT "Cloning and nucleotide sequence of the oriT region of the IncII plasmid R64."; RT J. Bacteriol. 170:4385-4387(1988). [10] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S typhimurium; PLASMID=R64; RA Furuya N., Nisioka T., Komano T.; RT "Nucleotide sequence and functions of the oriT operon in IncII plasmid R64."; RT J. Bacteriol. 173:2231-2237(1991). [11] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Furuya N., Nisioka T., Komano T.; RT "Determination of the nick site at oriT of IncII plasmid R64: global similarity of oriT structures of IncII and IncP plasmids."; RT J. Bacteriol. 175:5035-5042(1993). [12] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Kim S.R., Funayama N., Komano T.; RT "Nucleotide sequence and characterization of the traABCD region of IncII plasmid R64."; RT J. Bacteriol. 173:6612-6617(1991). [13] RN SEQUENCE FROM N.A. RP SPECIES=S. typhimurium; PLASMID=R64; RC SPECIES=S. typhimurium; PLASMID=R64; RA Furuya N., Komano T.; RT "Surface exclusion gene of IncII plasmid R64: nucleotide sequence and analysis of deletion mutants."; RT Plasmid 32:80-84(1994).

[14] SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
MEDLINE=96198148; PubMed=8626273;
RA Furuya N.; Komano T.;
RT "Nucleotide sequence and characterization of *trbABC* region of the Inc11 plasmid R64: existence and characterization of the *pnd* gene for plasmid maintenance within the transfer region.";
RT J. Bacteriol. 178:1491-1497 (1996).
RN [15] SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=97428559; PubMed=9281491;
RA Narahara K.; Rahman E.; Furuya N.; Komano T.;
RT "Requirement of a limited segment of the *sog* gene for plasmid R64 conjugation.";
RT Plasmid 38:1-11 (1997).
RN [16] SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=97315231; PubMed=9171405;
RA Kim S.R.; Komano T.;
RT "The Plasmid R64 thin pilus identified as a type IV pilus.";
RL J. Bacteriol. 179:3594-3603 (1997).
RN [17] SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=98053841; PubMed=9393692;
RA Furuya N.; Komano T.;
RT "Mutational analysis of the R64 *oriV* region: requirement for precise location of the NikaA-binding sequence.";
RL J. Bacteriol. 179:7291-7297 (1997).
RN [18] SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=98268996; PubMed=9603870;
RA Yoshida T.; Furuya N.; Ishikura M.; Isobe T.; Haino-Fukushima K.;
RA Ogawa T.; Komano T.;
RT "Purification and characterization of thin pili of Inc11 plasmids Colib-P9 and R64: formation of PilV-specific cell aggregates by type IV pili.";
RT J. Bacteriol. 180:2842-2848 (1998).
RN [19] SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=20223621; PubMed=10760136;
RA Komano T.; Yoshida T.; Narahara K.; Furuya N.;
RT "The transfer region of Inc11 plasmid R64: similarities between R64 tra genes and *Legionella* *icm/dot* genes.";
RL Mol. Microbiol. 35:1348-1359 (2000).
DR EMBL; J01829; AAA88660.1; -.
DR EMBL; AB026428; BAA83097.1; -.
DR EMBL; AF162223; AAD50250.1; -.
DR EMBL; AP000342; BAA78838.1; -.
DR InterPro; IPR002559; Transposase_1_1.
DR Pfam; PF01609; Transposase_1_1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 10
Q8GC95 PRELIMINARY; PRT; 402 AA.

QY 1 THHHP5 7
Db 247 THHHP5 253

[1] SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=01-MAR-2003 (TREMBLrel. 23, Created)
RA Furuya N.; Komano T.;
RT "Cloning, sequencing and expression of an invasion determinant from *Citrobacter freundii* strain 3009.";
RT Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AJ508060; CAD56977.1; -.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11
Q9H5N6 PRELIMINARY; PRT; 402 AA.
ID Q9H5N6
AC Q9H5N6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ23250.
OS Homo sapiens (Human).
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A.; Hikiji T.; Kobatake N.; Inagaki H.; Ikeda Y.; Okamoto S.;
RA Okitani R.; Ota T.; Suzuki Y.; Obayashi M.; Nishi T.; Shibahara T.;
RA Tanaka T.; Nakamura Y.; Isogai T.; Sugano S.;
RT "NEDO human cDNA sequencing project.";
RT Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AK026903; BAB15586.1; -.
DR InterPro; IPR002559; Transposase_1_1.
DR Pfam; PF01609; Transposase_1_1.
KW Hypothetical protein.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 4; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
Q9BTF6 PRELIMINARY; PRT; 402 AA.
ID Q9BTF6
AC Q9BTF6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A.; Hikiji T.; Kobatake N.; Inagaki H.; Ikeda Y.; Okamoto S.;
RA Okitani R.; Ota T.; Suzuki Y.; Obayashi M.; Nishi T.; Shibahara T.;
RA Tanaka T.; Nakamura Y.; Isogai T.; Sugano S.;
RT "NEDO human cDNA sequencing project.";
RT Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AK026903; BAB15586.1; -.
DR InterPro; IPR002559; Transposase_1_1.
DR Pfam; PF01609; Transposase_1_1.
KW Hypothetical protein.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 13
Q8GC95 PRELIMINARY; PRT; 402 AA.

QY 1 THHHP5 7
Db 247 THHHP5 253

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RC Champ M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC04131; AAH04131.1;
 DR InterPro; IPR002559; Transposase_11;
 DR Pfam; PF01609; Transposase_11;
 KW Hypothetical protein.
 SQ SEQUENCE 402 AA; 46041 MW; 2B4A8708C885667C CRC64;
 Query Match 88.6%; Score 39; DB 4; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Missmatches 1; Indels 0; Gaps 0;
 Qy 1 THIHHP 7
 Db 247 THCHHPS 253

RESULT 13

Q95TWO PRELIMINARY; PRT; 402 AA.
 ID Q95TWO;
 AC Q95TWO;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GM13045P.
 GN CG4800.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champ M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AY058482; AAL13711.1;
 DR FlyBase; FBgn0037874; CG4800.
 DR InterPro; IPR002559; Transposase_11;
 DR Pfam; PF01609; Transposase_11;
 SQ SEQUENCE 402 AA; 46041 MW; 2B4A8708C885667C CRC64;
 Query Match 88.6%; Score 39; DB 5; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Missmatches 1; Indels 0; Gaps 0;
 Qy 1 THIHHP 7
 Db 247 THCHHPS 253

RESULT 14

Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 19, Last sequence update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q8QTC0 PRELIMINARY; PRT; 402 AA.
 ID Q8QTC0
 AC Q8QTC0
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE WSSV410.
 OS White spot syndrome virus (WSSV).
 OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
 OX NCBI_TaxID=92652;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 MEDLINE=20517548; PubMed=11062040;
 RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
 RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
 RT "Identification and characterization of a shrimp white spot syndrome
 virus (WSSV) gene that encodes a novel chimeric polypeptide of
 RT cellular-type thymidine kinase and thymidylate kinase.";
 RT Virology 277:100-110(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 MEDLINE=21844071; PubMed=11853398;
 RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
 RA Lo C.F., Kou G.H.;
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
 RT spot syndrome virus and characterization of the motif important for
 RT targeting VP35 to the nuclei of transfected insect cells.";
 RL Targeting VP35 to the nuclei of transfected insect cells.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 RA Lo C.-F., Kou G.-H.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF440570; AAL89278.1;
 DR InterPro; IPR002559; Transposase_11;
 DR Pfam; PF01609; Transposase_11;
 SQ SEQUENCE 402 AA; 46073 MW; 2B4A8708D7930BAC CRC64;
 Query Match 88.6%; Score 39; DB 12; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Missmatches 1; Indels 0; Gaps 0;
 Qy 1 THIHHP 7
 Db 247 THCHHPS 253

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC Q95T96 PRELIMINARY; PRT; 402 AA.
 ID Q95T96;
 AC Q95T96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH10639P.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1] SEQUENCE

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds

(without alignments)

73.479 Million cell updates/sec

Title: US-09-901-187C-12

Perfect score: 34

Sequence: 1 MMMMMRL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rabbit:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriopl:*

17: sp_archeap:*

ALIGNMENTS

RESULT 1

Q9LZD2 ID PRELIMINARY; PRT; 70 AA.
 AC Q9LZD2; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 8.0 kDa protein.
 GN F12E4_330.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicots; Rosidae;
 OC euroids II; Brassicales; Arabidopsis; Arabidopsis;
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
 RA Lemcke K., Mayer K.F.X.; to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [12]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL SEQUENCE FROM N.A.
 RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
 RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Caminucci P.,
 RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RA "Full Length cDNA of gene F12E4_330 (GI:7378640)." ;
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RL DR AL162751; CAB83316_1;
 DR EMBL; AY040018; AAK64175_1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 70 AA; 8024 MW; A224B96F68370BC7 CRC64;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	32	94.1	70	Q9LZD2	Q9LZD2 arabidopsis	
2	30	88.2	192	Q9SQN9	Q9sqn9 arabidopsis	
3	30	88.2	249	Q8LDG1	Q8ldg1 arabidopsis	
4	30	88.2	249	Q9SVL0	Q9svl0 arabidopsis	
5	30	88.2	353	Q9STS7	Q9sts7 arabidopsis	
6	30	88.2	364	Q9AWZ8	Q9awz8 oryza sativ	
7	28	82.4	129	Q95RT0	Q95rt0 drosophila	
8	28	82.4	307	Q8WA93	Q8wa93 narceus ann	
9	28	82.4	307	Q8ZM25	Q8zm25 salmonella	
10	27	79.4	39	Q00868	Q00868 plasmodium	
11	27	79.4	70	Q8LE23	Q8le23 arabidopsis	
12	27	79.4	103	Q8H4V8	Q8h4v8 oryza sativ	
13	27	79.4	147	Q25870	Q25870 plasmodium	
14	27	79.4	176	Q8WBQ4	Q8wbq4 apis labori	
15	27	79.4	238	Q9FKE0	Q9fke0 arabidopsis	
16	27	79.4	283	Q8IQD9	Q8iqd9 drosophila	

Query Match 94.1%; Score 32; DB 10; Length 70;
 Best Local Similarity 85.7%; Pred. NO. 1.4;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMR 7
 Db 3 MMMMR 9

RESULT 2

Q9SQN9 PRELIMINARY; PRT; 192 AA.

ID Q9SQN9; PRELIMINARY; PRT; 249 AA.

AC Q9SQN9; ID Q9SQN9; PRELIMINARY; PRT; 249 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical protein.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;
 OX [1]

RN RP SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.

RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RA "Full-length messenger RNA sequences greatly improve genome
 annotation.";
 RL Genome Biol. 0:0-0 (2002).

RN [2]

RN RP SEQUENCE FROM N.A.

RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RA "Full-length cDNA from Arabidopsis thaliana.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AY086020; PAM63229.1;
 DR InterPro; IPR001356; Homeo_ZF_HD.

DR InterPro; IPR006455; Homeo_ZF_HD.

DR InterPro; IPR006456; ZF_HD_N.

DR TIGRFAMs; TIGR01565; homeo_ZF_HD_1.

DR TIGRFAMs; TIGR01566; ZF_HD_prot_N_1.

DR PROSITE; PS50071; HOMEobox_2; 1.

KW Hypothetical protein.

SQ SEQUENCE 249 AA; 28605 MW; 4CD673D6EEE665A7 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 249;
 Best Local Similarity 100.0%; Pred. NO. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMR 6
 Db 174 MMMMR 179

RESULT 4

Q9SVL0 PRELIMINARY; PRT; 249 AA.

ID Q9SVL0; PRELIMINARY; PRT; 249 AA.

AC Q9SVL0; ID Q9SVL0; PRELIMINARY; PRT; 249 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical 28.7 kDa protein.

GN P18B3_170.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;
 OX [1]

RN SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.

RA Cheuk R., Shinn P., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafai H., Bei Q., Chin C., Choi E., Johnson-Hopson C.,
 RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.R.;
 RA "Genomic sequence for Arabidopsis thaliana BAC F20B17 from chromosome
 I.";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RN SEQUENCE FROM N.A.

RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafai H., Bei B., Chin C., Choi E., Johnson-Hopson C.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.R.;
 RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC01171; AAG52257.1;
 DR EMBL; AC01093; AAF68112.1;
 DR HSSP; P29602; 1JER.

DR InterPro; IPR003245; Plcyanin_like.

DR Pfam; PF02298; Cu_bind_like; 1.

DR ProDom; PD003122; Plcyanin_like; 1.

KW Hypothetical protein.

DR Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,
 DR Mewes H.W., Mayer K.F.X., Lemcke K., Schueler C., Quettier F.,
 DR Salanoubat M.;
 DR Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.

RN EU Arabidopsis sequencing project;

RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN EMBL; AL049862; CAB42918.1;
 DR InterPro; IPR001356; Homeobox.

DR InterPro; IPR006455; Homeo_ZF_HD.

DR InterPro; IPR006456; ZF_HD_N.

DR InterPro; IPR006857; ZnF_HD_dimer.

Qy 1 MMMMR 6
 Db 2 MMMMR 7

RESULT 3

[1] Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY061159; AAL28707.1;
 DR FLYBase; FBgn0047212; BcDNA:LD12764;
 SEQUENCE 129 AA; 14240 MW; C349AAE9EC3DBEF CRC64;

Query Match Score 28; DB 5; Length 129;
 Best Local Similarity 85.7%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMMMMRL 7
 Db 13 MMMMTL 19

RESULT 8

Q8WA93 ID Q8WA93; PRELIMINARY; PRT; 307 AA.
 AC Q8WA93; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
 oxidoreductase chain 1).
 DE NAD1.
 OS Narceus annularius.
 OG Mitochondrion.
 OC Bukaryota; Metazoa; Arthropoda; Myriapoda; Diplopoda; Helmminthomorpha;
 Spirobolidae; Narceus.
 OC NCBITaxID=174156;

RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=21660208; PubMed=11801744;
 RX LAURO D.V.; Boore J.L.; Brown W.M.;
 "Complete mtDNA Sequences of Two Millipedes Suggest a New Model for
 Mitochondrial Gene Rearrangements: Duplication and Nonrandom Loss.";
 RT Mol. Biol. Evol. 19:163-169 (2002).
 RL [2] SEQUENCE FROM N.A.
 RP LAURO D.V.;
 RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -!- SIMILARITY: BELONGS TO THE COMPLEX 1 SUBUNIT 1 FAMILY.
 DR EMBL; AY055727; AAL18214.1;
 DR InterPro; IPR001694; Resp_NADH_dh1.
 DR Pfam; PF00146; NADHdh; 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 307 AA; 34669 MW; 03F837A142DBA24D CRC64;

Query Match Score 28; DB 8; Length 307;
 Best Local Similarity 85.7%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMMMMRL 7
 Db 233 MMMMSL 239

RESULT 9

Q8ZM25 ID Q8ZM25; PRELIMINARY; PRT; 307 AA.
 AC Q8ZM25; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative xylose/chitin deacetylase.
 GN STM3132.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OC NCBI_TaxID=602;

[1] SEQUENCE FROM N.A.
 RP STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RC MEDLINE=21534948; PubMed=11677609;
 RX McClelland M., Sanderson K.E.; Spieth J.J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete Genome sequence of *Salmonella enterica* serovar *Typhimurium*
 RT LT2.";
 RL Nature 413:852-856 (2001).
 DR EMBL; AE008844; AAL22006.1;
 DR InterPro; IPR002509; Polysac-deacet.
 DR Pfam; PF01522; Polysac-deacet; 1.
 KW Xylan degradation; Hydrolase; Glycosidase; Hypothetical protein;
 KW Complete proteome.
 SQ SEQUENCE 307 AA; 35194 MW; B14436094A935245 CRC64;

Query Match Score 28; DB 16; Length 307;
 Best Local Similarity 71.4%; Pred. No. 47;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMRL 7
 Db 249 MMMLRL 255

RESULT 10

Q00868 ID Q00868; PRELIMINARY; PRT; 39 AA.
 AC Q00868; 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE HMG-like protein PF16 (Fragment).
 GN PF16.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=5833;
 RN [1] SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB3;
 RX MEDLINE=92118046; PubMed=1731798;
 RA Guntaka R.V.; Kandala J.C.; Reddy V.D.;
 RT "Cloning and characterization of a highly conserved HMG-like protein
 (PF16) gene from Plasmodium falciparum.";
 RT (PF16) Biochem. Biophys. Res. Commun. 182:412-419 (1992).
 RL [2] SEQUENCE FROM N.A.
 RN [2] SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC Nambiar A.; Kandala J.C.; Dolan S.A.; Jensen J.J.; Guntaka R.V.;
 RA "Molecular cloning and characterization of a cDNA for the highly
 conserved HMG-like protein (PF16) gene of Plasmodium falciparum.";
 RT Submitted (APR 1997) to the EMBL/GenBank/DBJ databases.
 RL DR BMBL; U97128; AAB58116.1;
 RT NON_TER 39 39
 SQ SEQUENCE 39 AA; 5037 MW; 012D80DD1F566823 CRC64;

Query Match Score 27; DB 5; Length 39;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMRL 6
 Db 21 MMMMRK 26

RESULT 11

Q8LE23 ID Q8LE23; PRELIMINARY;
 AC Q8LE23; 01-OCT-2002 (TREMBLrel. 22, Created)

01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein.
 OS *Arabidopsis thaliana* (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TAXID=3702;
 RN [1] _SEQUENCE FROM N.A.
 RP Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.; "Full-Length cDNA from *Arabidopsis thaliana*.";
 RT Submitted (MAR-2002) to the EMBL/GenBank/DDJB databases.
 RL EMBL; AY085662; AAM67306.1; -.
 KW Hypothetical protein.
 SQ .SEQUENCE 70 AA; 7270 MW; 10C07644E0986E031 CRC64;
 [2] _SEQUENCE FROM N.A.
 RP Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.; "Full-Length messenger RNA sequences greatly improve genome
 annotation.";
 RT Annotation Biol. 0; 0-0 (2002).
 RL RN
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.; "Full-Length cDNA from *Arabidopsis thaliana*.";
 RT Submitted (MAR-2002) to the EMBL/GenBank/DDJB databases.
 RL EMBL; AY085662; AAM67306.1; -.
 KW Hypothetical protein.
 SQ .SEQUENCE 70 AA; 7270 MW; 10C07644E0986E031 CRC64;
 RESULT 12
 Q8H4V8 PRELIMINARY; PRT; 103 AA.
 ID Q8H4V8
 AC 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE QJ1484_G09.19 protein.
 GN QJ1484_G09.19
 OS *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 RN [1] _SEQUENCE FROM N.A.
 RP STRAIN=cv. Nipponbare;
 RC Sasaki T., Matsumoto T., Yamamoto K.;
 RA "Oryza sativa nippobare (GA33) genomic DNA, chromosome 8, BAC
 clone:QJ1484_G09.";
 RT Submitted (JUL-2001) to the EMBL/GenBank/DDJB databases.
 RL EMBL; AP003913; BAC24906.1; -.
 DR SEQUENCE 103 AA; 11753 MW; 22AB780BBF89PCB0 CRC64;
 SQ .SEQUENCE 79.4%; Score 27; DB 10; Length 103;
 Best Local Similarity 71.4%; Pred. No. 33;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MMMMMRL 7
 DB 45 MMVMMKL 51
 RESULT 13
 Q25870 PRELIMINARY; PRT; 147 AA.
 ID Q25870
 AC 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 RN [1] _SEQUENCE FROM N.A.
 RP MEDLINE=92118046; PubMed=1731798;
 RX Guntaka R.V., Kandala J.C., Vudem D.;
 RA "Cloning and characterization of a highly conserved HMG-like protein
 (EF16) gene from *Plasmodium falciparum*.";
 RT Biochem. Biophys. Res. Commun. 182:412-419 (1992).
 RL EMBL; M86518; AAA29615.1; -.
 DR EMBL; M86518; AAA29615.1; -.
 SQ .SEQUENCE 147 AA; 17175 MW; A46C5D47A62DF75A CRC64;
 RESULT 14
 Q8WBQ4 PRELIMINARY; PRT; 176 AA.
 ID Q8WBQ4
 AC 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE NADH dehydrogenase subunit 4 (Fragment).
 GN ND4.
 OS Apis laboriosa (Himalayan honeybee).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Apocrita; Aculeata; Apoidea;
 OC Apidae; Apis.
 NCBI_TAXID=183418;
 RN [1] _SEQUENCE FROM N.A.
 RA LeeJamanit W., Amano K.;
 RT "The NADH Dehydrogenases of *Apis mellifera*, *A. cerana*, *A. dorsata*, *A.
 florea* and *A. cerana* Sequence Comparison and Genetic Diversity.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DDJB databases.
 DR EMBL; AP447064; AAL60208.1; -.
 DR InterPro; IPR003918; NADHub_Oxred4.
 DR InterPro; IPR001750; Oxidored_q1.
 DR PFAM; PF00361; oxidored_q1; 1.
 DR PRINTS; PR01437; NUOXRDRTASE4.
 DR NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
 FW NON_TER 1 1
 FT NON_TER 176 176
 SQ SEQUENCE 176 AA; 19981 MW; 9E405EED9CF06687 CRC64;
 RESULT 15
 Q9FKE0 PRELIMINARY; PRT; 238 AA.
 ID Q9FKE0
 AC 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Gb|AAF30306.1.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBITaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Assamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and PAC clones.";
RT DNA Res. 5:203-216 (1998).
RL EMBL; AB012241; BAB09032.1;
DR InterPro; IPR006121; HeavyMe_transpC.
DR InterPro; IPR006191; Metal_bind.
DR Pfam; PF00403; HMA; 1.
DR PROSITE; PS50846; HMA; 2; 1.
SEQUENCE 238 AA; 27717 MW; F2F486C355011632 CRC64;
Query Match 79.4%; Score 27; DB 10; Length 238;
Best Local Similarity 71.4%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MMMMMRL 7
Db 103 MMMMM 109

Search completed: February 11, 2004, 17:09:48
Job time : 26.5833 SECS

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
 (without alignments)
 63.714 Million cell updates/sec

Title: US-09-901-187C-12
 Perfect score: 34
 Sequence: 1 MMMMMRL 7

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	85.3	132	1 F802 SCHMA	P16464 schistosoma
2	27	79.4	146	1 2SS_BEREX	P04403 bertholleti
3	27	79.4	1610	1 CCAD_MESEAU	Q99244 mesocricetus
4	27	79.4	2161	1 CCAD_HUMAN	Q01668 homo sapien
5	27	79.4	2203	1 CCAD_RAT	P27732 rattus norvegicus
6	26	76.5	395	1 RIF2_YEAST	Q06208 saccharomyces
7	26	76.5	403	1 HRPN_ERWAM	Q01099 erwinia amylovora
8	26	76.5	438	1 POP1_CAELA	Q10666 caenorhabditis
9	26	76.5	875	1 AMD2_XENLA	P12890 xenopus laevis
10	25	73.5	313	1 NULM_RHISA	Q99824 rhinocerophalus
11	25	73.5	321	1 AOX1_SOYBN	Q07185 glycine max
12	25	73.5	325	1 CY1_HUMAN	P08574 homo sapien
13	25	73.5	355	1 HKL6_LYCSES	Q22299 lycopersicon esculentum
14	25	73.5	377	1 Y4WD_RHISN	P55682 rhizobium
15	25	73.5	382	1 STM_BRACOL	Q38874 arabidopsis thaliana
16	25	73.5	383	1 STM_BRACOL	Q9m6d9 brassica oleracea
17	25	73.5	400	1 AMD1_XENLA	P08478 xenopus laevis
18	25	73.5	457	1 NORM_ECOLI	P58164 escherichia coli
19	25	73.5	554	1 NUSM_APILI	P37340 escherichia coli
20	25	73.5	554	1 YKR2_CAEEL	P34308 caenorhabditis elegans
21	25	73.5	783	1 DING_BACSU	P54394 bacillus subtilis
22	25	73.5	931	1 CLV3_ARATH	Q9y691 homo sapiens
23	24	70.6	96	1 PD11_HUMAN	P56983 apis mellifera
24	24	70.6	99	1 Y281_METUJ	Q57729 methanococcus
25	24	70.6	179	1 GC12_PSEAE	Q9i351 pseudomonas
26	24	70.6	181	1 GC11_PSEAE	Q9hy98 pseudomonas
27	24	70.6	186	1 CTRC_NEIMA	P57012 neisseria meningitidis
28	24	70.6	265	1 PYRB_XYLFA	Q9pbb8 xylella fastidiosa
29	24	70.6	322	1 MBI2_YEAST	P03873 saccharomyces cerevisiae
30	24	70.6	423	1 AVT_CATCO	Q90352 catostomus
31	24	70.6	434	1 PUR2_XYLFA	Q9pc09 xylella fastidiosa
32	24	70.6	437	1	

ALIGNMENTS

Q9ape0_klebsiella	34	24	70.6	462	1 ZRAS_KLEBOX
Q16850_homo_sapien	35	24	70.6	503	1 CP51_HUMAN
Q8zyu6_pyrobaculum	36	24	70.6	562	1 ILYD_PYRAE
Q11046_mycobacteri	37	24	70.6	582	1 YC73_MYCTU
P32569_saccharomyces	38	24	70.6	687	1 SRB4_YEAST
P93024_arabidopsis	39	24	70.6	902	1 MOPT_ARATH
P53599_saccharomyces	40	24	70.6	1579	1 SSK2_YEAST
OB3508_trponema	41	23	67.6	49	1 Y495_TREPA
P45708_bacillus_suis	42	23	67.6	72	1 YNEF_BACSU
Q8csp3_staphylococ	43	23	67.6	75	1 YA26_STAEP
Q99ud3_staphylococcus	44	23	67.6	80	1 YD43_STAAM
Q29555_archaeoglob	45	23	67.6	89	1 Y703_ARCFU

RESULT 1
 F802 SCHMA STANDARD; PRT; 132 AA.

ID P802 SCHMA STANDARD; PRT; 132 AA.
 AC P16464;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Female specific 800 protein (FS800).
 OS Schistosoma mansoni (Blood fluke).
 OC Burkaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeida; Schistosomatidae; Schistosomatidae; Schistosoma mansoni.";

RT NCBITaxID=6183;

RN [11]

RP SEQUENCE FROM N.A.

RC STRAIN_Puerto Rican;

RA MEDLINE=89181810; PubMed=2927441;

RT "Localization and pattern of expression of a female specific mRNA in

RT Schistosoma mansoni".

RL Mol. Biochem. Parasitol. 32:113-119 (1989).

CC -!- FUNCTION: FS800 is likely to have some function in the production or maintenance of the schistosome egg. It may have a function unrelated to eggshell formation.

CC -!- DEVELOPMENTAL STAGE: Highest level only in mature worms, i.e., during egg production.

CC -!- MISCELLANEOUS: The two F800 proteins are read from two overlapping reading frames.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

CC DR J03999; AAA29884.1; - ILE-RICH.
 FT DOMAIN 2 21 POLY-MET.
 FT DOMAIN 47 53 POLY-MET.

FT SEQUENCE 15561 MW; 02C77F42A25E120E CRC64;

Query Match 85.3%; Score 29; DB 1; Length 132;
 Best Local Similarity 71.4%; Pred. No. 2;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMRL 7
 Db 49 MMMMMKI 55

RESULT 2
 2SS_BEREX STANDARD; PRT; 146 AA.
 ID 2SS_BEREX
 AC P04403; P04402;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)
 DT 2S sulfur-rich seed storage protein precursor (Allergen Ber e 1).
 DE BE2S1 AND BE2S2.
 OS Bertholletia excelsa (Brazil nut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Ericales; Lecythidaceae; Bertholletia.
 OC NCBITaxID=3645;
 OX [1]
 RN SEQUENCE FROM N.A.
 RA Altenbach S.B., Pearson K.W., Leung F.W., Sun S.S.M.;
 RT "Cloning and sequence analysis of a cDNA encoding a Brazil nut protein
 exceptionally rich in methionine.";
 RL Plant Mol. Biol. 8:239-250(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bassunner R.; Submitted (DEC-1990) to the EMBL/GenBank/DDJB databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91370890; PubMed=1840683;
 RA Gander E.S., Holmstrom K.O., de Paiva G.R., de Castro L.A.B.,
 RA Carneiro M., Grossi de Sa M.F.;
 RT "Isolation, characterization and expression of a gene coding for a 2S
 albumin from Bertholletia excelsa (Brazil nut).";
 RT Plant Mol. Biol. 16:437-448(1991).
 RN [4]
 RP SEQUENCE OF 37-64 AND 70-142.
 RX MEDLINE=87004679; PubMed=3758080;
 RA Ampe C., van Damme J., de Castro L.A.B., Sampaio M.J.A.M.,
 RA van Montagu M., Vandekerckhove J.;
 RT "The amino-acid sequence of the 2S sulphur-rich proteins from seeds
 of Brazil nut (Bertholletia excelsa H.B.K.).";
 RL Eur. J. Biochem. 159:597-604(1986).
 CC -!- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -!- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -!- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M17146; AAA33010.1;
 DR EMBL; X57027; CAA40343.1;
 DR EMBL; X57028; CAA40344.1;
 DR EMBL; X54490; CAA38362.1;
 DR EMBL; X54491; CAA38363.1; ALT_SEQ.
 DR EMBL; A13818; CAA01131.1; -.
 DR PIR; A25802; A25802.
 DR PIR; S14946; S14946.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR000617; Napin.
 DR Pfam; PF00234; tryp_alpha_amyl; 1.
 DR PRINTS; PR00496; NAPIN.
 DR SMART; SM00499; AAI; 1.
 KW 3D-structure.
 FT SIGNAL 1 22
 FT PROPEP 23 36
 FT CHAIN 37 64
 FT PROPEP 65 69
 FT CHAIN 70 142
 FT PROPEP 143 146
 FT MOD RES 37 37
 FT VARIANT 91 91
 FT CONFLICT 38 39
 FT CONFLICT 122 122
 FT CONFLICT 126 126
 SQ 146 AA; 16911 MW; A7DF778FD766410D CRC64;
 Query Match 79.4%; Score 27; DB 1; Length 146;
 Best Local Similarity 83.3%; Pred. No. 6.7;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MMMMRL 7
 DB 99 MMMMRL 104
 RESULT 3
 CCAD_MESAU STANDARD;
 ID CCAD_MESAU
 AC 099244; Q99245;
 AC 099244; Q99245;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Voltage-dependent L-type calcium channel alpha-1D subunit (Calcium
 channel, L type, alpha-1 polypeptide isoform 2).
 DE CACN1A1 OR CACN1A2 OR CACN1A2 OR CACN3 OR CACN4.
 GN Mesocricetus auratus (Golden hamster).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM HCA3A).
 RC TISSUE=Insulinoma;
 RX MEDLINE=93149124; PubMed=1337146;
 RA Yaney G.C., Wheeler M.B., Wei X., Perez-Reyes E., Birnbaumer L.,
 RA Boyd A.E. III, Moss L.G.;
 RA "Cloning of a novel alpha 1-subunit of the voltage-dependent calcium
 channel from the beta-cell.";
 RL Mol. Endocrinol. 6:2143-2152 (1992).
 RN [2]
 RP SEQUENCE OF 1146-1441 FROM N.A. (ISOFORMS CACH3B AND CACH3D).
 RC TISSUE=Heart;
 RX MEDLINE=91056091; PubMed=2173707;
 RA Perez-Reyes E., Wei X., Castellano A., Birnbaumer L.;
 RT "Molecular diversity of L-type calcium channels. Evidence for
 alternative splicing of the transcripts of three non-allelic genes.";
 RL J. Biol. Chem. 265:20430-20436 (1990).
 CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
 CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
 CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
 CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
 CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
 CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
 CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
 CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
 CC PHENYLALKYLAMINES, BENZOTHIAZEPINES, AND BY OMEGA-AGATOXIN-111A
 CC (OMEGA-AGA-111A). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
 CC GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-1VA (OMEGA-AGA-1VA).
 CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE FORE-
 CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
 CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Additional isoforms seem to exist;
 CC Name=RCA3A;
 CC IsoId=Q99244-1; Sequence=Displayed;
 CC Name=CACH3B;
 CC IsoId=Q99244-3; Sequence=Not described;
 CC Name=CACH3D;
 CC IsoId=Q99244-2; Sequence=VSP_000915;
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART AND SKELETAL MUSCLE.

-!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.

-!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHAI-1 SUBUNITS FAMILY.

CC EMBL; M57969; AAB59702; 1; - .
 CC EMBL; M57970; AAA62807; 1; - .
 CC PIR; A46227; A46227.
 CC DR InterPro; IPR001682; Ca/Na_pore.
 CC DR InterPro; IPR002077; Ca_channel.
 CC DR InterPro; IPR002111; Cat_channel_TrpL.
 CC DR InterPro; IPR005821; Ion_trans.
 CC DR InterPro; IPR005446; LVDCALpha1.
 CC DR InterPro; IPR005452; LVDCALpha1D.
 CC DR InterPro; IPR005820; M+channel_nlg.
 CC DR PF00520; ion_trans; 4.
 CC DR PRINTS; PR00167; C CHANNEL.
 CC DR PRINTS; PR01630; LVDCALPHA1.
 DR PRO1636; LVDCALPHA1D.
 DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;
 KW Calcium-binding; Glycosylation; Phosphorylation; Alternative splicing.
 FT REPEAT 112 408 I.
 FT REPEAT 508 754 II.
 FT REPEAT 871 1153 III.
 FT REPEAT 1190 1465 IV.
 FT DOMAIN 1 125 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 126 144 S1 OF REPEAT I (POTENTIAL).
 FT DOMAIN 145 162 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 163 182 S2 OF REPEAT I (POTENTIAL).
 FT DOMAIN 183 194 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 195 213 S3 OF REPEAT I (POTENTIAL).
 FT DOMAIN 214 234 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 235 253 S4 OF REPEAT I (POTENTIAL).
 FT DOMAIN 254 272 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 273 292 S5 OF REPEAT I (POTENTIAL).
 FT DOMAIN 293 380 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 381 405 S6 OF REPEAT I (POTENTIAL).
 FT TRANSMEM 523 542 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 406 522 S1 OF REPEAT II (POTENTIAL).
 FT TRANSMEM 523 543 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 543 557 S2 OF REPEAT II (POTENTIAL).
 FT TRANSMEM 558 576 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 577 584 S3 OF REPEAT II (POTENTIAL).
 FT TRANSMEM 585 603 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 604 613 S4 OF REPEAT II (POTENTIAL).
 FT TRANSMEM 614 632 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 633 651 S5 OF REPEAT II (POTENTIAL).
 FT TRANSMEM 652 672 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 673 726 S6 OF REPEAT II (POTENTIAL).
 FT TRANSMEM 727 751 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 752 884 S1 OF REPEAT III (POTENTIAL).
 FT TRANSMEM 885 903 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 904 919 S2 OF REPEAT III (POTENTIAL).
 FT TRANSMEM 920 939 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 940 951 S3 OF REPEAT III (POTENTIAL).
 FT TRANSMEM 952 970 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 971 976 S4 OF REPEAT III (POTENTIAL).
 FT TRANSMEM 977 996 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 997 1015 S5 OF REPEAT III (POTENTIAL).
 FT TRANSMEM 1016 1035 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1036 1125 S6 OF REPEAT III (POTENTIAL).
 FT TRANSMEM 1126 1146

CC CYTOPLASMIC (POTENTIAL).
 CC S1 OF REPEAT IV (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC S2 OF REPEAT IV (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC S3 OF REPEAT IV (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC S4 OF REPEAT IV (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC S5 OF REPEAT IV (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC S6 OF REPEAT IV (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC POLY-MET.
 CC POLY-LEU.
 CC POLY-GLU.
 CC BINDING TO THE BETA SUBUNIT
 CC (BY SIMILARITY).
 CC CALCIUM ION SELECTIVITY AND PERMEABILITY
 CC (BY SIMILARITY).
 CC CALCIUM ION SELECTIVITY AND PERMEABILITY
 CC (BY SIMILARITY).
 CC CALCIUM ION SELECTIVITY AND PERMEABILITY
 CC (BY SIMILARITY).
 CC CALCIUM ION SELECTIVITY AND PERMEABILITY
 CC (BY SIMILARITY).
 CC DIHYDROPYRIDINES (BY SIMILARITY).
 CC DIHYDROPYRIDINES (BY SIMILARITY).
 CC PHENYLALKYLAMINES (BY SIMILARITY).
 CC PHOSPHORYLATION (BY PKA) (POTENTIAL).
 CC POTENTIAL.
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC GYFSDAWNTEDSLIVIGSIIIDVALSEADPTESESSLPLPTAT
 PG -> HYFTDAWNTEDSLIVIGSIIIDVALSEADPTESESSLPLPTAT
 Isoform CACH3D).
 /FTId=VSP 000915.

FT SITE 363 363
 FT SITE 704 704
 FT SITE 1099 1099
 FT SITE 1404 1404
 FT BINDING 1073 1163
 FT BINDING 1418 1484
 FT BINDING 1430 1473
 FT MOD_RES 1473 1502
 FT CA_BIND 1491 154
 FT CARBOHYD 224 224
 FT CARBOHYD 328 328
 FT VARSPLIC 1261 1303
 FT FT
 FT SEQUENCE 1610 AA; 182327 MN; B3B2E3794D936F79 CRC64;
 FT SEQUENCE 1610 AA; 182327 MN; B3B2E3794D936F79 CRC64;

Query Match 79.4%; Score 27; DB 1; Length 1610;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMMAL 6
 Db 2 MAMMALK 7

RESULT 4
 CCAD_HUMAN ID_CCAD_HUMAN STANDARD; PRT; 2161 AA.
 AC Q01668; Q13916; Q13931;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Voltage-dependent L-type calcium channel alpha-1D subunit (Calcium
 channel, L type, alpha-1 polypeptide, isoform 2).
 DE CACNA1D OR CACNL1A2 OR CACNA1D OR CACNA1A2 OR CACNA1D
 GN Homo sapiens (Human)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1] RP SEQUENCE FROM N.A. (ISOFORM NEURONAL-TYPE).
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=92110010; PubMed=1309651;
 RA Williams M.E., Feldman D.H., McCue A.F., Brenner R., Velicelebi G.,
 RA Ellis S.B., Harpold M.M.;
 RT "Structure and functional expression of alpha 1, alpha 2, and beta
 RT subunits of a novel human neuronal calcium channel subtype.";
 RL Neuron 8:71-84 (1992).
 RN [2]

SEQUENCE FROM N.A. (ISOFORM BETA-CELL-TYPE).
 TISSUE=Pancreatic islets;
 MEDLINE=92115705; PubMed=1309948;
 Seino S., Chen L., Seino M., Blondel O., Takeda J., Johnson J.H.,
 Bell G.I.; "Cloning of the alpha 1 subunit of a voltage-dependent calcium
 channel expressed in pancreatic beta cells";
 Proc. Natl. Acad. Sci. U.S.A. 89:584-588(1992).
 [3]

SEQUENCE FROM N.A. (ISOFORM BETA-CELL-TYPE), AND VARIANT NIDDM.
 MEDLINE=96044438; PubMed=7557998;
 RX Yamada Y., Masuda K., Li Q., Ihara Y., Kubota A., Miura T.,
 RA Nakamura K., Fujii Y., Seino S., Seino Y.; "The structures of the human calcium channel alpha 1 subunit
 (CACNL1A2) and beta subunit (CACNLB3) genes";
 Genomics 27:312-319 (1995).
 -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
 ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
 IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
 CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
 CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
 GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
 CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
 GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
 PHENYLALKYLAMINES, BENZOTIAZEPINES, AND BY OMEGA-AGATOXIN-III
 (OMEGA-AGA-III). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
 GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
 -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
 FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
 SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Comment=Additional isoforms seem to exist;

CC Name=Neuronal-type;
 CC IsoId=Q01668-1; Sequence=Displayed;
 CC Name=Beta-cell-type;
 CC IsoId=Q01668-2; Sequence=vsp_000913; vsp_000914;
 CC TISSUE SPECIFICITY: EXPRESSED IN PANCREATIC ISLETS AND IN BRAIN,
 WHERE IT HAS BEEN SEEN IN HIPPOCAMPUS, BASAL GANGLIA, HABENULA AND
 THALAMUS. NO EXPRESSION IN SKELETAL MUSCLE.
 -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1', S2', S3', S5', S6') AND ONE
 PROBABLY CHARGED TRANSMEMBRANE SEGMENT (S4'). S4 SEGMENTS
 PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 -!- POLYMORPHISM: A CHANGE FROM SEVEN TO EIGHT ATG TRINUCLEOTIDE
 REPEATS, RESULTING IN AN ADDITIONAL N-TERMINAL METHIONINE, HAS
 BEEN FOUND IN A PATIENT WITH NON-INSULIN-DEPENDENT DIABETES
 MELLITUS (NIDDM).
 -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

CC EMBL; M76558; AAA58402.1; -
 DR EMBL; M83566; AAA35629.1; -
 DR EMBL; D43747; BAA07804.1; -
 DR EMBL; D43705; BAA07804.1; JOINED.
 DR EMBL; D43706; BAA07804.1; JOINED.
 DR EMBL; D43707; BAA07804.1; JOINED.
 DR EMBL; D43708; BAA07804.1; JOINED.
 DR EMBL; D43709; BAA07804.1; JOINED.

CC EMBL; D43710; BAA07804.1; JOINED.
 DR EMBL; D43711; BAA07804.1; JOINED.
 DR EMBL; D43712; BAA07804.1; JOINED.
 DR EMBL; D43713; BAA07804.1; JOINED.
 DR EMBL; D43714; BAA07804.1; JOINED.
 DR EMBL; D43715; BAA07804.1; JOINED.
 DR EMBL; D43716; BAA07804.1; JOINED.
 DR EMBL; D43717; BAA07804.1; JOINED.
 DR EMBL; D43718; BAA07804.1; JOINED.
 DR EMBL; D43719; BAA07804.1; JOINED.
 DR EMBL; D43720; BAA07804.1; JOINED.
 DR EMBL; D43721; BAA07804.1; JOINED.
 DR EMBL; D43722; BAA07804.1; JOINED.
 DR EMBL; D43723; BAA07804.1; JOINED.
 DR EMBL; D43724; BAA07804.1; JOINED.
 DR EMBL; D43725; BAA07804.1; JOINED.
 DR EMBL; D43726; BAA07804.1; JOINED.
 DR EMBL; D43727; BAA07804.1; JOINED.
 DR EMBL; D43728; BAA07804.1; JOINED.
 DR EMBL; D43729; BAA07804.1; JOINED.
 DR EMBL; D43730; BAA07804.1; JOINED.
 DR EMBL; D43731; BAA07804.1; JOINED.
 DR EMBL; D43732; BAA07804.1; JOINED.
 DR EMBL; D43733; BAA07804.1; JOINED.
 DR EMBL; D43734; BAA07804.1; JOINED.
 DR EMBL; D43735; BAA07804.1; JOINED.
 DR EMBL; D43736; BAA07804.1; JOINED.
 DR EMBL; D43737; BAA07804.1; JOINED.
 DR EMBL; D43738; BAA07804.1; JOINED.
 DR EMBL; D43739; BAA07804.1; JOINED.
 DR EMBL; D43740; BAA07804.1; JOINED.
 DR EMBL; D43741; BAA07804.1; JOINED.
 DR EMBL; D43742; BAA07804.1; JOINED.
 DR EMBL; D43743; BAA07804.1; JOINED.
 DR EMBL; D43744; BAA07804.1; JOINED.
 DR EMBL; D43745; BAA07804.1; JOINED.
 DR EMBL; D43746; BAA07804.1; JOINED.
 DR Genew; HGNC:1391; CACNA1D.
 DR MIM; 114206; -
 DR GO; GO:0005891; C:voltage-gated calcium channel complex; TAS.
 DR GO; GO:0015270; F:dihydropyridine-sensitive calcium channel a...; TAS.
 DR GO; GO:0006832; P:small molecule transport; TAS.
 DR InterPro; IPR002177; Ca_Channel.
 DR InterPro; IPR002111; Cat_Channel_TPL.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR005446; LVDCCALpha1.
 DR InterPro; IPR005452; LVDCCALpha1D.
 DR InterPro; IPR005820; M+channel_nlg.
 DR Pfam; PF00520; ion_trans; 4.
 DR Prints; PR00167; CACHANNEL.
 DR Prints; PR01630; LVDCCALPHA1.
 DR Prints; PR01636; LVDCCALPHA1D.
 DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;
 KW Calcium-binding; Phosphorylation; Alternative splicing; Polymorphism;
 KW Triplet repeat expansion.

CC FT REPEAT 113 409 I.
 CC FT REPEAT 509 755 II.
 CC FT REPEAT 873 1155 III.
 CC FT REPEAT 1192 1467 IV.
 CC FT DOMAIN 1 126 S1 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 127 145 S1 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 146 163 EXTRACELLULAR (POTENTIAL).
 CC FT DOMAIN 164 183 S2 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 184 195 S3 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 196 214 S3 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 215 235 S4 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 236 254 S4 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 255 273 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 274 293 S5 OF REPEAT I (POTENTIAL).
 CC FT DOMAIN 294 381 EXTRACELLULAR (POTENTIAL).
 CC FT DOMAIN 382 406 S6 OF REPEAT I (POTENTIAL).

CC IsoId=P27732-3; Sequence=vsp_000921;
 CC Name=4; Synonyms=Delta-IV-S3;
 CC IsoId=P27732-4; Sequence=vsp_000919;
 CC Name=5; Synonyms=Delta-IV-S4;
 CC IsoId=P27732-5; Sequence=vsp_000922;
 CC Name=6; Synonyms=RB9;
 CC IsoId=P27732-6; Sequence=vsp_000920, vsp_000921;
 CC Name=7; Synonyms=RB11;
 CC IsoId=P27732-7; Sequence=vsp_000917;
 CC Name=8; Synonyms=RB34;
 CC IsoId=P27732-8; Sequence=vsp_000916;
 CC Name=9; Synonyms=RH1;
 CC IsoId=P27732-9; Sequence=vsp_000918;
 CC Name=10; Synonyms=RH2;
 CC IsoId=P27732-10; Sequence=vsp_000919, vsp_000922;
 CC Name=11; Synonyms=RKC5;
 CC IsoId=P27732-11; Sequence=Not described;
 CC Name=12; Synonyms=RKC6;
 CC IsoId=P27732-11; Sequence=vsp_000919;
 CC Name=13; Synonyms=ROB3;
 CC IsoId=P27732-14; Sequence=Not described;
 CC Name=14; Synonyms=Truncated;
 CC IsoId=P27732-12; Sequence=vsp_000925, vsp_000926;
 CC TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PANCREATIC ISLETS AND B-
 CC LYMPHOCYTES.
 CC DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR D38101; BAA07282.1; -
 DR EMBL; D38102; BAA07283.1; -
 DR EMBL; M57682; AAA42015.1; -
 DR EMBL; U14005; AAB60515.1; -
 DR EMBL; M99221; AAA40895.1; -
 DR EMBL; U31772; AAA89156.1; -
 DR EMBL; U49126; AAB61634.1; -
 DR EMBL; U49127; AAB61635.1; -
 DR EMBL; U49128; AAB61636.1; -
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR002077; Ca_channel.
 DR InterPro; IPR02111; Cat_channel_TriPL.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR003091; K_channel.
 DR InterPro; IPR005446; LVDCCALpha1D.
 DR InterPro; IPR005452; LVDCCALpha1D.
 DR InterPro; IPR005820; M+channel_nlg.
 DR Pfam; PF00520; ion_trans; 4.
 DR PRINTS; PRO0167; C CHANNEL.
 DR PRINTS; PRO0169; KC CHANNEL.
 DR PRINTS; PR01630; LVDCCALPHA1.
 DR PRINTS; PR01636; LVDCCALPHA1D.
 KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;
 KW Calcium-binding; Phosphorylation; Alternative splicing.

PT REPEAT 112 408 I.
 PT REPEAT 528 774 II.
 PT REPEAT 892 1174 III.
 PT REPEAT 1211 1486 IV.
 PT DOMAIN 1 126 CYTOPLASMIC (POTENTIAL).
 PT TRANSMEM 127 145 S1 OF REPEAT I (POTENTIAL).
 PT DOMAIN 146 163 EXTRACELLULAR (POTENTIAL).
 CC -----
 RESULT 6
 RIF2 YEAST
 ID RIF2 YEAST
 AC Q06208;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE RIF2 Protein (RAPI-interacting factor 2).
 GN RIF2 OR YLR453C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBi_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=9169871; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
 RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
 RA Entian K.-D., Floeth M., Goffeau A., Heblig U., Heumann K.,
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
 RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
 RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urrestarazu L.A., Vandembol M., Verhasselt P.,
 RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
 RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RT Nature 387:87-90(1997).
 RN [2]
 RP FUNCTION, AND SUBUNIT.
 RX MEDLINE=97242430; PubMed=9087429;
 RA Wotton D., Shore D.;
 RT "A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to
 RT regulate telomere length in Saccharomyces cerevisiae.";
 RT Genes Dev. 11:748-760(1997).
 CC -----
 -!- FUNCTION: Involved in transcriptional silencing and telomere

length regulation. Its role in telomere length regulation results from either a block in elongation or promoting degradation of the telomere ends. Loss of RIF1 function results in derepression of an HMR silencer, whose ARS consensus element has been deleted, and in the elongation of telomeres. RAP1 may target the binding of RIF1 to silencers and telomeres.

-!- SUBUNIT: Interacts with RIF1 and RAP1 C-terminus.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U22382; AAB67535.1; -.
 DR PIR; S55975; S55975.
 DR SGD; S0004445; RIF2.
 DR GO; GO:0005696; C:telomere; IDA.
 DR GO; GO:0042162; F:telomeric DNA binding activity; IDA.
 DR GO; GO:007004; P:telomerase-dependent telomere maintenance; IMP.
 KW Telomere.
 SQ SEQUENCE 395 AA;
 FF6E3A22C8805DE9 CRC64;

Query Match 76.5%; Score 26; DB 1; Length 395;
 Best Local Similarity 83.3%; Pred. No. 30;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M~~MM~~MR 6
 191 M~~MM~~MR 196
 Db

RESULT 7
 HRPN_ERWAM STANDARD; PRT; 403 AA.
 AC Q01099;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Harpin (Harpin-EA).
 GN HRPN.
 OS Erwinia amylovora.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Erwinia.
 OX NCBI_TaxID=552;
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
 RC STRAIN=Ec321;
 RX MEDLINE=92320301; PubMed=1621099;
 RA Wei Z.-M., Laby R.J., Zumoff C.H., Bauer D.W., He S.Y., Collier A.,
 RA Beer S.V.;
 RA "Harpin, elicitor of the hypersensitive response produced by the
 plant pathogen *Erwinia amylovora*.";
 RT Science 257:85-88(1992).
 RL [2]

RN REVISIONS.
 RC STRAIN=Ec321;
 RA Laby R.J., Kim J.F., Beer S.V.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ELICTICS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT
 CC UPON INFECTION. HARPIN ELICTICS HR IN NON-HOSTS AND IS ALSO
 CC REQUIRED FOR PATHOGENICITY IN HOST PLANTS.
 CC -!- SUBCELLULAR LOCATION: SECRETED; VIA THE HRP SECRETION PATHWAY.

CC [2]
 RN REVISIONS.
 RC STRAIN=Ec321;
 RA Laby R.J., Kim J.F., Beer S.V.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ELICTICS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT
 CC UPON INFECTION. HARPIN ELICTICS HR IN NON-HOSTS AND IS ALSO
 CC REQUIRED FOR PATHOGENICITY IN HOST PLANTS.
 CC -!- SUBCELLULAR LOCATION: SECRETED; VIA THE HRP SECRETION PATHWAY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U37532; AAC05308.1; -.

CC HSSP; P27782; 2LEF.

CC TRANSFAC; T03244; -.

CC InterPro; IPR000910; HMG_12_box.

CC Pfam; PF00050; HMG_box; 1.

CC SMART; SM00398; HMG_1.

CC PROSITE; PS50118; HMG_BOX_2; 1.

KW DNA-binding; Nuclear protein; Developmental protein.

FT DOMAIN 55 173 PRO-RICH.

FT DOMAIN 132 139 POLY-ALA.

FT DOMAIN 192 262 HMG BOX.

FT DOMAIN 340 347 POLY-SER.

FT DOMAIN 359 362 POLY-GLN.

FT DOMAIN 411 416 POLY-GLU.

SQ SEQUENCE 438 AA; 48595 MW; F94073BE40B4095F CRC64;

Query Match 76.5%; Score 26; DB 1; Length 438;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;
Sequence 1 MMMMMRL 7 369 MLMQMRL 375

RESULT 9
AMD2_XENLA STANDARD PRT; 875 AA.

ID P12B50;

AC 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Peptidyl-glycine alpha-amidating monooxygenase II precursor

DE (EC 1.14.17.3) (Peptide C-terminal alpha-amidating enzyme II) (AE-II).

OS Xenopus laevis (African Clawed Frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;

OC Xenopodinae; Xenopus.

NCBI_TaxID=8355; [1]

RN SEQUENCE FROM N.A.

RC TISSUE-Skin; PMID=88134244; PubMed=2829895;

RA Ohsuye K., Kitano K., Wada Y., Fuchimura K., Tanaka S., Mizuno K.,

RA Matsuo H.,

RT "Cloning of cDNA encoding a new peptide C-terminal alpha-amidating enzyme having a putative membrane-spanning domain from *Xenopus laevis* skin".

RT Biochem. Biophys. Res. Commun. 150:1275-1281(1988)

CC -!- FUNCTION: C-terminal alpha-amidating monooxygenase II precursor. The reaction produces a peptidyl(2-hydroxyglycine) intermediate is unstable and dismutates to glyoxylate and the corresponding desglycine peptide amide.

CC -!- CATALYTIC ACTIVITY: Peptidylglycine + ascorbate + O(2) = Peptidyl(2-hydroxyglycine) + dehydroascorbate + H(2)O.

CC -!- COFACTOR: COPPER AND ASCORBATE.

CC -!- SIMILARITY: BELONGS TO THE COPPER TYPE II, ASCORBATE-DEPENDENT MONOOXYGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/monooxygenase.html> or send an email to license@isb-sib.ch).

CC EMBL; AF081829; AAD05524.1; -.

CC PIR; T11160; T11160.

CC InterPro; IPR001694; Resp_NADH_dh1.

CC DR PROSITE; PS00667; COMPLEX1_ND1_1;

CC DR PROSITE; PS00667; COMPLEX1_ND1_2;

CC DR KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.

CC DR SEQUENCE 313 AA; 36721 MW; 846DA5BFB56SEAD1 CRC64;

Query Match 76.5%; Score 26; DB 1; Length 875;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Sequence 1 MMMMMRL 7 358 MMMMMRM 364

Query Match 76.5%; Score 26; DB 1; Length 875;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Sequence 1 MMMMMRL 7 358 MMMMMRM 364

RESULT 10
NU1M_RHISA STANDARD PRT; 313 AA.

ID NU1M_RHISA STANDARD PRT; 313 AA.

AC 099824; [1]

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).

GN NDL.

OS Rhinicephalus sanguineus (Brown dog tick).

OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

OC Parasitiformes; Ixodida; Ixodidae; Rhinicephalus.

NCBI_TaxID=34632; [1]

RN SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RX MEDLINE=99083443; PubMed=9866211;

RA Black W.C. IV, Roehrdanz R.L.; [1]

RT "Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes.";

RT Mol. Biol. Evol. 15:1772-1785(1998).

RL -!- CATALYTIC ACTIVITY: NADH + Ubiquinone = NAD(+) + ubiquinol.

CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

CC RN [1]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/complex1.html> or send an email to license@isb-sib.ch).

CC CC EMBL; AF081829; AAD05524.1; -.

CC CC PIR; T11160; T11160.

CC CC InterPro; IPR001694; Resp_NADH_dh1.

CC CC DR PROSITE; PS00667; COMPLEX1_ND1_1;

CC CC DR PROSITE; PS00667; COMPLEX1_ND1_2;

CC CC DR KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.

CC CC DR SEQUENCE 313 AA; 36721 MW; 846DA5BFB56SEAD1 CRC64;

CC CC DR PROSITE; PS00084; CU2_monoxygense.

CC CC DR PROSITE; PS00085; CU2_monoxygense.

CC CC DR PROSITE; PS00084; CU2_monoxygense.

CC CC DR PROSITE; PS00085; CU2_monoxygense.

CC CC DR KW Oxidoreductase; Monoxygenase; Copper; Vitamin C; Transmembrane;

RESULT 11	AOX1_SOYBN	STANDARD;	PRT;	321 AA.	Db	1 M00MM 5
AC Q07186; Q41265;						
DT 01-NOV-1997 (Rel. 35, Created)						
DT 01-NOV-1997 (Rel. 35, Last sequence update)						
DT 28-FEB-2003 (Rel. 41, Last annotation update)						
DE Alternative oxidase 1, mitochondrial precursor (EC 1.1.1.1).						
GN AOX1.						
OS Glycine max (Soybean).						
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;						
OC euroids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.						
NCBI_TaxID=3847;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Shoot; MEDLINE=94120028; PubMed=8290651;						
RX RA Whelan J.M., McIntosh L., Day D.A.; "Sequencing of a soybean alternative oxidase cDNA clone."; Plant Physiol. 103:1481-1481(1993).						
RT RL [2]						
RN RP SEQUENCE OF 194-233 FROM N.A. MEDLINE=96165778; PubMed=8580775;						
RX RA Whelan J., Millar A.H., Day D.A.; "The alternative oxidase is encoded in a multigene family in soybean."; Plant 198:197-201(1996).						
CC CC -!- FUNCTION: CATALYZES CYANIDE-RESISTANT OXYGEN CONSUMPTION. MAY RESTRICTED, OR IN RESPONSE TO LOW TEMPERATURES (BY SIMILARITY).						
CC CC -!- PATHWAY: Alternative respiratory pathway.						
CC CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (PROBABLE).						
CC CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL, POSSIBLY IN THE INNER SURFACE OF THE INNER MITOCHONDRIAL MEMBRANE.						
CC CC -!- SIMILARITY: BY salicylic acid.						
CC CC -!- INDUCTION: BELONGS TO THE ALTERNATIVE OXIDASE FAMILY.						
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).						
CC DR X6B702; CAA48653.1; -.						
DR EMBL; S81466; AAC34192.1; -.						
DR InterPro; IPR002680; AOX.						
DR Pfam; PF01786; AOX; 1.						
KW Oxidoreductase; Trans peptide; Mitochondrion; Respiratory chain; Inner membrane; Transmembrane; Multigene family.						
FT FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).						
FT CHAIN ? ALTERNATIVE OXIDASE 1.						
FT TRANSMEM 146 166 POTENTIAL.						
FT TRANSMEM 208 228 POTENTIAL.						
FT DISULFID 94 94 INTERCHAIN (POTENTIAL).						
FT METAL 112 112 POTENTIAL.						
FT METAL 165 165 POTENTIAL.						
FT METAL 192 192 POTENTIAL.						
FT METAL 233 233 POTENTIAL.						
FT METAL 294 294 POTENTIAL.						
FT METAL 299 299 POTENTIAL.						
FT DOMAIN 1 5 POLY-MET.						
SQ SEQUENCE 321 AA; 36437 MW; E00FF981F0C2D0E7 CRC64;						
Query Match 73.5%; Score 25; DB 1; Length 321;						
Best Local Similarity 100.0%; Pred. No. 43;						
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
CC 1 M00MM 5						

CC B-C1 COMPLEX, WHICH ACCEPTS ELECTRONS FROM RIESKE PROTEIN AND
 CC TRANSFERS ELECTRONS TO CYTOCHROME C IN THE MITOCHONDRIAL
 CC RESPIRATORY CHAIN.
 CC -!- SUBUNIT: B-C1 COMPLEX CONTAINS 11 SUBUNITS; 3 RESPIRATORY SUBUNITS
 CC (CYTOCHROME B, CYTOCHROME C1, RIESKE PROTEIN), 2 CORE PROTEINS AND
 CC 6 LOW-MOLECULAR WEIGHT PROTEINS.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial intermembrane space.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC DR EMBL; J04444; AAA52135.1; -.
 CC DR EMBL; M16597; AAA35730.1; -.
 CC DR EMBL; BC001006; AAH01006.1; -.
 CC DR EMBL; BC015616; AAH15616.1; -.
 CC DR EMBL; BC020566; AAH20566.1; -.
 CC DR EMBL; X06994; CAA30052.1; -.
 CC DR PIR; A31481; S00680;
 CC GeneW; HGNC; 2579; CYC1.
 CC DR MIM; 123980; -.
 CC DR GO; GO:0005739; C:mitochondrion; TAS.
 CC DR GO; GO:0009461; F:cytochrome c; TAS.
 CC DR InterPro; IPR002326; Cyt_C1.
 CC DR InterPro; IPR000345; Cyt_C_heme_bind.
 CC DR Pfam; PF02167; Cytochrome_C1; 1.
 CC DR PRINTS; PRO0603; CYTOCHROME_C1.
 CC DR PROSITE; PS00190; CYTOCHROME_C; 1.
 CC KW Electron transport; Respiratory chain; Heme; Mitochondrion;
 CC KW Transmembrane; Transm. peptide; Polymorphism.
 CC PT TRANSIT 1 84 MITOCHONDRIAL.
 CC FT CHAIN 85 325 CYTOCHROME C1, HEME PROTEIN.
 CC FT BINDING 121 121 HEME (COVALENT).
 CC FT BINDING 124 124 HEME (COVALENT).
 CC FT METAL 244 244 IRON (HEME AXIAL LIGAND).
 CC FT METAL 245 245 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC FT TRANSMEM 292 306 ANCHORS TO THE MEMBRANE (POTENTIAL).
 CC FT VARIANT 89 89 L -> V.
 CC FT SEQUENCE 325 AA; 35390 MW; CC8815E60E99EBDC CRC64;
 CC FT FTID=VAR_013631.

CC Query Match 73.5%; Score 25; DB 1; Length 325;
 CC Best Local Similarity 71.4%; Pred. No. 43;
 CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC Qy 1 MNNNNRL 7
 CC Db 292 MNNNNMAL 298

CC Query Match 73.5%; Score 25; DB 1; Length 355;
 CC Best Local Similarity 70.0%; Pred. No. 47;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC Qy 1 MNNNN 5
 CC Db 47 MNNNN 51

RESULT 13
 ID HKL6_LYCES STANDARD; PRT; 355 AA.
 AC 022299;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein knotted-1 like LET6.
 GN LET6.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Lamiidae; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=CV_VENT_Cherry;
 RX MEDLINE=98145476; PubMed=9484482;
 RA Janssen B.J., Williams A., Chen J.J., Mathern J., Hake S., Sinha N.;
 RP SEQUENCE FROM N.A.
 OC Rhizobiaceae; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Plasmid sym PNGR234a.
 OC Rhizobium sp. (strain NGR234).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=97305956; PubMed=9163424;
 RX

RT "Isolation and characterization of two knotted-like homeobox genes
 RT from tomato.";
 RT Plant Mol. Biol. 36:417-425 (1998).
 RL CC -!- FUNCTION: MAY HAVE A ROLE TO PLAY IN FORMATIVE EVENTS IN OVULE AND
 CC EMBRYO MORPHOGENESIS. PROBABLY BINDS TO THE DNA SEQUENCE 5'-TGAC-
 CC 3'.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING LATERAL ORGANS AND
 CC DEVELOPING OVARIES IN FLOWERS.
 CC -!- SIMILARITY: BELONGS TO THE TALE/HOMEOBOX FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF000141; AAC49917.1; -.
 CC DR PIR; T04317; T04317.
 CC DR HSSP; P40424; 1B72.
 CC DR TRANSFAC; T04057; -.
 CC DR InterPro; IPR005539; ELK.
 CC DR InterPro; IPR001356; Homeobox.
 CC DR InterPro; IPR005540; KNOX1.
 CC DR InterPro; IPR0005541; KNOX2.
 CC DR Pfam; PF03789; ELK; 1.
 CC DR Pfam; PF00046; homeobox; 1.
 CC DR Pfam; PF03790; KNOX1; 1.
 CC DR Pfam; PF03791; KNOX2; 1.
 CC DR PRODOM; PD000010; Homeobox; 1.
 CC DR SMART; SM00389; HOX; 1.
 CC DR PROSITE; PS000027; HOMEobox_1; 1.
 CC DR PROSITE; PS50071; HOMEBOX_2; 1.
 CC KW Homeobox; DNA-binding; Nuclear protein.
 CC FT DOMAIN 24 33 POLY-ASN.
 CC FT DOMAIN 47 51 POLY-MET.
 CC FT DOMAIN 52 56 POLY-PRO.
 CC FT DOMAIN 60 63 POLY-ASN.
 CC FT DOMAIN 79 84 POLY-ASN.
 CC FT DOMAIN 91 95 POLY-SER.
 CC FT DOMAIN 142 146 POLY-SER.
 CC FT DOMAIN 235 258 ELK DOMAIN.
 CC FT DNA BIND 259 321 HOMEBOX (TALE-TYPE).
 CC SQ SEQUENCE 355 AA; 39796 MW; SESALAFD75808C49 CRC64;
 CC Query Match 73.5%; Score 25; DB 1; Length 355;
 CC Best Local Similarity 100.0%; Pred. No. 47;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC Qy 1 MNNNN 5
 CC Db 47 MNNNN 51

RESULT 14
 ID Y4WD_RHISN STANDARD; PRT; 377 AA.
 AC P55682;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical transport protein Y4WD.
 GN Y4WD.
 OS Rhizobium sp. (strain NGR234).
 OC Rhizobiaceae; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Plasmid sym PNGR234a.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=CV_VENT_Cherry;
 RX MEDLINE=98145476; PubMed=9484482;
 RA Janssen B.J., Williams A., Chen J.J., Mathern J., Hake S., Sinha N.;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;

Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A., "Molecular basis of symbiosis between Rhizobium and legumes."; Perre X.; Nature 387:394-401(1997).

RT -!- FUNCTION: COULD BE INVOLVED IN A TRANSPORT SYSTEM.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -!- SIMILARITY: TO R.MELIOTI MOSC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR AB00103; AAB91911.1; InterPro; IPR007114; MFS.

CC KW Hypothetical protein; Transmembrane; Transport; Plasmid.

FT TRANSMEM 4 24 POTENTIAL.

FT TRANSMEM 41 61 POTENTIAL.

FT TRANSMEM 85 105 POTENTIAL.

FT TRANSMEM 134 154 POTENTIAL.

FT TRANSMEM 159 179 POTENTIAL.

FT TRANSMEM 192 212 POTENTIAL.

FT TRANSMEM 278 298 POTENTIAL.

FT TRANSMEM 301 321 POTENTIAL.

FT TRANSMEM 327 347 POTENTIAL.

FT TRANSMEM 356 376 POTENTIAL.

SQ SEQUENCE 377 AA; 39051 MW; 49CF6E44AA0D74BD CRC64;

Query Match 73.5%; Score 25; DB 1; Length 377;

Best Local Similarity 83.3%; Pred. No. 50;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 M0MMLR 7

Db 250 M0MMLR 255

RESULT 15

STM_ARATH ID STM_ARATH STANDARD; PRT; 382 AA.

AC Q38874; Q8RXJ1; Q9MAV3;

DT 15-JUL-1998 (Rel. 36, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Homeobox protein SHOOT MERISTEMLESS.

GN STM OR AT1G62360 OR F2401.9.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_TaxID=3702;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=CV. Wassilewskija;

RX MEDLINE=96135134; PubMed=8538741;

RA Long J.A., Moan E.I., Medford J.I., Barton M.K.; "A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis.";

RT Nature 379:66-69(1996).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=CV. Columbia;

RX MEDLINE=21016719; PubMed=11130712;

RA Theologis A., Ecker J.R., Palm C.J., Feder Spiel N.A., Kaul S., Altafai H., Araujo R., Bowman C.L., Brooks S.Y., White C., Alonso J., Buehler E., Chan A., Chao Q., Conn L., Conway A.B., Conway T.H., Creasy T.H., Dewar K., Chung M.K., Dunn P., Etgui P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.P., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,

RA Kim C.J., Koo H.I., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luos J.S., Maiti R., Marziali A., Millitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambungga G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.; "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."; Nature 408:816-820 (2000).

RN [3]

RP SEQUENCE OF 57-382 FROM N.A.

RC STRAIN=CV. Columbia;

RA Shinnozaki K., Davis R.W., Ecker J.R., Theologis A.; "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the SSP consortium (Salk/Stanford/BGEC)." ; Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.

RL [4]

RN FUNCTION.

RP MEDLINE=21932426; PubMed=11934861;

RX Byrne M.E., Simorowski J., Martienssen R.A.; "ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis." ; Development 129:1957-1965(2002).

RL -!- FUNCTION: Required for shoot apical meristem formation during embryogenesis. Negatively regulates ASYMMETRIC LEAVES1 (AS1) and ASYMMETRIC LEAVES2 (AS2 or LBD6). Probably binds to the DNA sequence 5'-TGAC-3'.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- TISSUE SPECIFICITY: Expressed in all four types of shoot apical meristems (SAM) i.e., in vegetative, axillary, inflorescence and floral.

CC -!- DEVELOPMENTAL STAGE: First expressed in early to mid-globular-stage embryos. In late globular stage, detected as a stripe running medially across the top of the embryo. In heart stage embryo, expression is restricted to a notch between the cotyledons. In seedlings and adult plants found in all shoot apical meristems. In the inflorescence meristem, expression disappears as floral buds are initiated and reappears in the later floral meristem where it is found in the central portion of the developing gynoecium. Also detected in the L1 layer of embryo.

CC -!- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEobox FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/> or send an email to license@isb-sib.ch).

CC EMBL; U32344; AAC49148.1; -.

DR EMBL; AC003113; AAF70849.1; -.

DR PIR; S68456; S68456.

DR PIR; T01446; T01446.

DR HSSP; P41778; 1DU6.

DR InterPro; IPR005539; ELK.

DR InterPro; IPR001356; Homeobox.

DR InterPro; IPR005540; KNOX1.

DR InterPro; IPR005541; KNOX2.

DR Pfam; PF03789; ELK; 1.

DR Pfam; PF00046; homeobox; 1.

DR Pfam; PF03790; KNOX1; 1.

DR Pfam; PF03791; KNOX2; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEobox-1; 1.

DR PROSITE; PS50071; HOMEobox-2; 1.

KW DNA-binding; Homeobox; NucTear protein; Developmental protein.

FT DOMAIN 283 ELK DOMAIN

FT DNA_BIND 284 346 HOMEobox (TALE-TYPE).

FT DOMAIN 26 3 0 POLY-MET.
FT DOMAIN 37 6 0 HIS-RICH.
FT DOMAIN 83 8 9 POLY-SER.
FT DOMAIN 113 12 0 POLY-ALA.
FT DOMAIN 161 16 6 S -> F (IN REF. 1).
FT CONFLICT 115 115 MISSING (IN REF. 2).
FT CONFLICT 229 23 3 G -> D (IN REF. 1).
FT CONFLICT 369 36 9 SQ 382 AA; 42753 MW; 6227D3DB4093E732 CRC64;
SEQUENCE 382 AA; 42753 MW; 6227D3DB4093E732 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNNNN 5
Db 26 MNNNN 30

Search completed: February 11, 2004, 17:04:17
Job time : 6.16667 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-12

Perfect score: 34

Sequence: 1 MNMDMRL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : PIR 76:*

1: Pir1;*
2: Pir2;*
3: Pir3;*
4: Pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	70	2 T48381	hypothetical prote
2	30	88.2	192	2 H96828	hypothetical prote
3	30	88.2	249	2 T08410	hypothetical prote
4	30	88.2	353	2 T07725	hypothetical prote
5	27	79.4	39	2 PC4294	high mobility grou
6	27	79.4	146	2 S14946	2S seed storage pr
7	27	79.4	147	2 JQ1320	high mobility grou
8	27	79.4	427	2 T04869	transforming prote
9	27	79.4	1610	2 A46227	voltage-dependent
10	27	79.4	1646	2 JH0422	calcium channel al
11	27	79.4	2161	2 JH0564	calcium channel al
12	27	79.4	2181	2 A38198	voltage-dependent
13	27	79.4	2203	2 T42742	DNA-3-methyladenin
14	26	76.5	216	2 H82695	protein F3F9.20 [1
15	26	76.5	302	2 H96811	mutants block spor
16	26	76.5	309	2 B83999	protein W03D8.4 [1
17	26	76.5	340	2 C87732	hypothetical prote
18	26	76.5	340	2 T32931	hypothetical prote
19	26	76.5	375	2 T06096	hypothetical prote
20	26	76.5	395	2 S55975	hypothetical prote
21	26	76.5	403	2 T08471	harpin - Erwinia a
22	26	76.5	436	2 E96635	hypothetical prote
23	26	76.5	438	1 A57667	pop-1 protein - Ca
24	26	76.5	484	2 H95377	probable Trk-like
25	26	76.5	875	1 URXLA2	peptidylglycine mo
26	25	73.5	68	2 T03645	hypothetical prote
27	25	73.5	98	2 A98222	hypothetical prote
28	25	73.5	120	2 AI2181	probable regulator
29	25	73.5	124	2 F81942	

ALIGNMENTS

30	25	73.5	124	2 A81163
31	25	73.5	135	2 F86239
32	25	73.5	176	2 A86159
33	25	73.5	197	2 T26156
34	25	73.5	217	2 H86657
35	25	73.5	220	2 AD0225
36	25	73.5	242	2 B82815
				hypothetical prote
				myb-related trans
				hypothetical prote
				probable amino-aci
				conserved hypothet
				protein F20B24.4 [
				hypothetical prote
				hypothetical prote
				probable amino-aci
				conserved hypothet
				protein F20B24.4 [
				hypothetical prote
				NADH2 dehydrogenas
				alternative respir
				ubiquinol-cytochro
				chitinase (EC 3.2.

37	25	73.5	268	2 T04787
38	25	73.5	275	2 T51651
39	25	73.5	289	2 T29216
40	25	73.5	292	2 T0930
41	25	73.5	307	2 AD0883
42	25	73.5	313	2 T11160
43	25	73.5	321	2 S31711
44	25	73.5	325	1 S00680
45	25	73.5	327	2 T09687

30	25	73.5	124	2 A81163
31	25	73.5	135	2 F86239
32	25	73.5	176	2 A86159
33	25	73.5	197	2 T26156
34	25	73.5	217	2 H86657
35	25	73.5	220	2 AD0225
36	25	73.5	242	2 B82815
				hypothetical prote
				myb-related trans
				hypothetical prote
				probable amino-aci
				conserved hypothet
				protein F20B24.4 [
				hypothetical prote
				NADH2 dehydrogenas
				alternative respir
				ubiquinol-cytochro
				chitinase (EC 3.2.

37	25	73.5	124	2 A81163
38	25	73.5	135	2 F86239
39	25	73.5	176	2 A86159
40	25	73.5	197	2 T26156
41	25	73.5	217	2 H86657
42	25	73.5	220	2 AD0225
43	25	73.5	242	2 B82815
44	25	73.5	268	2 T04787
45	25	73.5	275	2 T51651

30	25	73.5	124	2 A81163
31	25	73.5	135	2 F86239
32	25	73.5	176	2 A86159
33	25	73.5	197	2 T26156
34	25	73.5	217	2 H86657
35	25	73.5	220	2 AD0225
36	25	73.5	242	2 B82815
37	25	73.5	268	2 T04787
38	25	73.5	275	2 T51651
39	25	73.5	289	2 T29216
40	25	73.5	292	2 T0930
41	25	73.5	307	2 AD0883
42	25	73.5	313	2 T11160
43	25	73.5	321	2 S31711
44	25	73.5	325	1 S00680
45	25	73.5	327	2 T09687

30</td

A;Map position: 1
 C;Superfamily: stellacyanin

Query Match 88.2%; Score 30; DB 2; Length 192;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 6; Conservative 0; Missmatches 0; Indels 0; Gaps 0;

Oy 1 M|MMMR 6
 Db 2 M|MMMR 7

RESULT 3
 T08410 hypothetical protein F18B3.170 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
 C;Accession: T08410
 R;Quétier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salamouti
 Submitted to the Protein Sequence Database, May 1999
 A;Reference number: Z16409
 A;Accession: T08410
 A;Molecule type: DNA
 A;Residues: 1-249 <QUE>
 A;Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.170
 A;Experimental source: cultivar Columbia; BAC clone F18B3
 C;Genetics:
 A;Gene: ATSP:F18B3.170
 A;Map position: 3

Query Match 88.2%; Score 30; DB 2; Length 249;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 6; Conservative 0; Missmatches 0; Indels 0; Gaps 0;

Oy 1 M|MMMR 6
 Db 174 M|MMMR 179

RESULT 4
 T07725 hypothetical protein T23J7.190 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 22-Oct-1999
 C;Accession: T07725
 R;Quétier, F.; Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigou
 Submitted to the Protein Sequence Database, April 1999
 A;Reference number: Z15793
 A;Accession: T07725
 A;Molecule type: DNA
 A;Residues: 1-353 <QUE>
 A;Cross-references: EMBL:AL049746; GSPDB:GN00061; ATSP:T23J7.190
 A;Experimental source: cultivar Columbia; BAC clone T23J7
 C;Genetics:
 A;Gene: ATSP:T23J7.190
 A;Map position: 3
 A;Introns: 28/2; 83/3; 165/3; 262/3

Query Match 88.2%; Score 30; DB 2; Length 353;
 Best Local Similarity 100.0%; Pred. No. 7.6;
 Matches 6; Conservative 0; Missmatches 0; Indels 0; Gaps 0;

Oy 1 M|MMMR 6
 Db 122 M|MMMR 127

RESULT 5
 PC4294 high mobility group protein Pf16 - malaria parasite (*Plasmodium falciparum*) (fragments)
 C;Species: Plasmodium falciparum
 C;Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jun-2000
 C;Accession: PC4294

A;Map position: 1
 C;Superfamily: wheat alpha-amylase inhibitor

R;Nambiar, A.; Kandala, J.C.; Dolan, S.A.; Jensen, J.B.; Guntaka, R.V.
 Biochem. Biophys. Res. Commun. 234, 101-106, 1997
 A;Title: Molecular cloning and characterization of a cDNA for the highly conserved HMG-
 A;Reference number: JCS403; MUID:97312528; PMID:9168969
 A;Accession: PC4294
 A;Molecule type: mRNA
 A;Residues: 1-39 <NAME>
 C;Comment: This protein is involved in DNA replication, chromatin assembly and transcri-
 C;Genetics:
 A;Gene: Pf16

Query Match 79.4%; Score 27; DB 2; Length 39;
 Best Local Similarity 83.3%; Pred. No. 5.2;
 Matches 5; Conservative 1; Missmatches 0; Indels 0; Gaps 0;

Oy 1 M|MMMR 6
 Db 21 M|MMMR 26

RESULT 6
 S14946 2S seed storage protein large chain - Brazil nut
 N;Alternate names: albumin 2S precursor
 C;Species: Bertholletia excelsa (Brazil nut)
 C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
 C;Accession: S14946; S14479; S06252; S21640; B25802
 R;Gander, E.S.; Holmstrom, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Gross
 Plant Mol. Biol. 16, 437-448, 1991
 A;Title: Isolation, characterization and expression of a gene coding for a 2S albumin E
 A;Reference number: S14946; MUID:91370890; PMID:1840683
 A;Accession: S14946
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-146 <GAN>
 A;Cross-references: EMBL:X54490; NID:917710; PID:CAA38362.1; PID:917711
 A;Note: the authors translated the codon CTT for residue 13 as Val and GTC for residue
 R;Bassuner, R.; Schlesier, B.
 Submitted to the EMBL Data Library, December 1990
 A;Reference number: S14479
 A;Accession: S14479
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-146 <BAS>
 A;Cross-references: EMBL:X57027; NID:917714; PID:917715
 R;Altenbach, S.B.; Pearson, K.W.; Leung, F.W.; Sun, S.S.M.
 Plant Mol. Biol. 8, 239-250, 1987
 A;Title: Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptic
 A;Reference number: S06252
 A;Accession: S06252
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-146 <ALT>
 A;Experimental source: clone PHS-3
 A;Note: Part of this sequence, including the amino end of the large chain, was confirmed
 R;Bassuner, R.; Schlesier, B.
 Submitted to the EMBL Data Library, December 1990
 A;Reference number: S21640
 A;Accession: S21640
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-146 <BA2>
 A;Cross-references: EMBL:X57028; NID:917716; PID:917717
 R;Ampe, C.; Van Damme, J.; de Castro, L.A.B.; Sampaio, M.J.A.M.; Van Montagu, M.; Vand
 Eur. J. Biochem. 159, 597-604, 1986
 A;Title: The amino-acid sequence of the 2S sulphur-rich proteins from seeds of Brazil
 A;Reference number: A91173; MUID:87004679; PMID:3758080
 A;Accession: B25802
 A;Molecule type: Protein
 A;Residues: 70-90, 'E', '92-121, 'M', 123-125, 'L', 127-142 <AMP>
 C;Genetics:
 A;Introns: 60/3
 C;Superfamily: wheat alpha-amylase inhibitor

F;1-22/Domain: signal sequence #status predicted <SIG>
 F;23-46/Domain: propeptide #status predicted <PRO>
 F;47-69/Product: seed storage protein small chain #status predicted <SCH>
 F;70-146/Product: seed storage protein large chain #status experimental <LCH>

Query Match	79.4%;	Score 27;	DB 2;	Length 146;	
Best Local Similarity	83.3%;	Pred. No. 18;			
Matches 5;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	2 MNNMRL 7	;			
Db	99 MNNM RM 104				

RESULT 7
 JQ1320
 high mobility group protein Pf16 - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jun-2000

C;Accession: JQ1320

R;Guntaka, R.V.; Kandala, J.C.; Reddy, V.D.

Biochem. Biophys. Res. Commun. 182, 412-419, 1992

A;Title: Cloning and characterization of a highly conserved HMG-like protein (PF16) gene

A;Reference number: JQ1320; MUID:92118046; PMID:1731798

A;Accession: JQ1320

A;Molecule type: DNA

A;Residues: 1-147 <GTUN>

A;Cross-references: GB:M86518; MJD:g160325; PID:g160326

C;Comment: This protein interacts with other nuclear proteins and serve as a transcripti

C;Genetics:

A;Gene: PF16

F;7-18,19-30/Region: duplication

F;70-91/Region: aspartic acid/glutamic acid-rich

F;126-133/Region: basic

Query Match

79.4%;

Score 27;

DB 2;

Length 147;

Best Local Similarity

83.3%;

Pred. No. 18;

Matches 5;

Conservative 1;

Mismatches 0;

Indels 0;

Gaps 0;

Qy

1 MNNMRL 6

|||||;

Db

11 MNNM RM 16

RESULT 8
 T04869
 transforming protein myb homolog F28A21.180 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 19-May-2000

C;Accession: T04869

R;Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer, R.; Submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15387

A;Accession: T04869

A;Molecule type: DNA

A;Residues: 1-427 <BEV>

A;Cross-references: EMBL:AL035526

A;Experimental source: Cultivar Columbia; BAC clone F28A21

C;Genetics:

A;Map position: 4

A;Introns: 226/2; 268/3

A;Note: F28A21.180

C;Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolo

F;212-263/Domain: myb DNA-binding repeat homology <MYB>

Query Match

79.4%;

Score 27;

DB 2;

Length 427;

Best Local Similarity

83.3%;

Pred. No. 47;

Matches 5;

Conservative 1;

Mismatches 0;

Indels 0;

Gaps 0;

Qy

2 MNNMRL 7

|||||;

Db

145 MNNM RM 150

RESULT 9
 A46227
 voltage-dependent Ca2+ channel alpha 1-subunit - golden hamster

C;Species: Mesocricetus auratus (golden hamster)

C;Date: 27-Oct-1993 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000

C;Accession: A46227

R;Yaney, G.C.; Wheeler, M.B.; Wei, X.; Perez-Reyes, E.; Birnbaumer, L.; Boyd III, A.E.;

Mol. Endocrinol. 6, 2143-2152, 1992

A;Title: Cloning of a novel alpha 1-subunit of the voltage-dependent calcium channel for

A;Reference number: A46227; MUID:93149124; PMID:1337146

A;Accession: A46227

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1610 <YAN>

A;Experimental source: insulin-secreting cell line HIT-T15

A;Note: sequence extracted from NCBI backbone (NCBIP:123692)

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match

79.4%;

Score 27;

DB 2;

Length 1610;

Best Local Similarity

83.3%;

Pred. No. 1.6e+02;

Matches 5;

Conservative 1;

Mismatches 0;

Indels 0;

Gaps 0;

Qy

1 MNNMRL 6

|||||;

Db

2 MNNM RM 7

|||||;

Db

145 MNNM RM 150

RESULT 10
 JH0422
 voltage-dependent calcium channel complex alpha-1 chain - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jan-2000

C;Accession: JH0422; D35901; T60901

R;Hui, A.; Ellinor, P.T.; Krizanova, O.; Wang, J.J.; Diebold, R.J.; Schwartz, A.

Neuron 7, 3544, 1991

A;Title: Molecular cloning of multiple subtypes of a novel rat brain isoform of the alpha

A;Reference number: JH0422; MUID:91299338; PMID:1646940

A;Accession: JH0422

A;Molecule type: mRNA

A;Residues: 1-1646 <HUI>

A;Cross-references: GB:M57682; PID:9206573; PID:AA42015.1; PID:9206574

A;Experimental source: brain

R;Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.; Davidson, N.

Proc. Natl. Acad. Sci. U.S.A. 87, 3391-3395, 1990

A;Title: Rat brain expresses a heterogeneous family of calcium channels.

A;Reference number: A35901

A;Accession: D35901

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual transcript

A;Molecule type: mRNA

A;Residues: 1247-1434 <SNU>

A;Experimental source: brain

R;Yu, A.S.L.; Hebert, S.C.; Brenner, B.M.; Lytton, J.

Proc. Natl. Acad. Sci. U.S.A. 89, 10494-10498, 1992

A;Title: Molecular characterization and nephron distribution of a family of transcripts

A;Reference number: A46422; MUID:93066265; PMID:1279681

A;Accession: I60901

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1040-1261,1305-1365 <RES>

A;Cross-references: GB:M99221; PID:9203370; PID:AAA40895.1; PID:9203371

A;Experimental source: kidney

C;Comment: Calcium channels are essential for many cellular functions, such as muscle c

n.

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

C;Keywords: alternative splicing; calcium binding; glyccoprotein; phosph

F;1463-1491/Domain: calcium binding #status predicted <EPC>

F;154,224,328/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;464,848,1489,1584/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match

79.4%;

Score 27;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMR 6
Db 2 MMMMMK 7

RESULT 11

JH0564 calcium channel alpha-1D chain - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 24-Nov-1999
C;Accession: JH0564
R;Williams, M.E.; Feldman, D.H.; McCue, A.F.; Brenner, R.; Velicelebi, G.; Ellis, S.B.; Neuron 8, 71-84, 1992
A;Reference number: JH0564; MUID:92110010; PMID:1309651
A;Accession: JH0564
A;Molecule type: mRNA
A;Residues: 1-2161 <WIL>
A;Cross-references: GB: M76558
A;Experimental source: neuroblastoma, cell line IMR32
C;Comment: This protein is a subunit of the voltage-dependent calcium channel.
C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C;Keywords: transmembrane protein
F;127-145/Domain: transmembrane #status predicted <IS1>
F;164-183/Domain: transmembrane #status predicted <IS2>
F;195-214/Domain: transmembrane #status predicted <IS3>
F;236-254/Domain: transmembrane #status predicted <IS4>
F;274-293/Domain: transmembrane #status predicted <IS5>
F;382-406/Domain: transmembrane #status predicted <IS6>
F;524-542/Domain: transmembrane #status predicted <II1>
F;558-577/Domain: transmembrane #status predicted <II2>
F;586-603/Domain: transmembrane #status predicted <II3>
F;615-633/Domain: transmembrane #status predicted <II4>
F;653-672/Domain: transmembrane #status predicted <II5>
F;728-752/Domain: transmembrane #status predicted <II6>
F;888-905/Domain: transmembrane #status predicted <S1I>
F;922-941/Domain: transmembrane #status predicted <S2I>
F;954-972/Domain: transmembrane #status predicted <S3I>
F;980-998/Domain: transmembrane #status predicted <S4I>
F;1018-1037/Domain: transmembrane #status predicted <S5I>
F;1128-1152/Domain: transmembrane #status predicted <S6I>
F;1206-1224/Domain: transmembrane #status predicted <VS1>
F;1240-1259/Domain: transmembrane #status predicted <VS2>
F;1268-1286/Domain: transmembrane #status predicted <VS3>
F;1315-1333/Domain: transmembrane #status predicted <VS4>
F;1353-1372/Domain: transmembrane #status predicted <VS5>
F;1440-1464/Domain: transmembrane #status predicted <VS6>

RESULT 13

T42742 voltage-dependent calcium channel alpha 1 chain, isoform CACN4A - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T42742
R;Ihara, Y.; Yamada, Y.; Fujii, Y.; Gonori, T.; Yano, H.; Yasuda, K.; Inagaki, N.; Seir Mol. Endocrinol. 9, 121-130, 1995
A;Title: Molecular diversity and functional characterization of voltage-dependent calcium channels
A;Reference number: 222258; MUID:95280950; PMID:7760845
A;Accession: T42742
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2203 <WHA>
A;Cross-references: EMBL:D38101; NID:9736711; PIDN:BAA07282.1; PID:9736712
A;Experimental source: insulinoma RNase complementary DNA library
C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C;Keywords: calcium channel

Query Match 79.4%; Score 27; DB 2; Length 2203;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMR 6
Db 3 MMMMMK 8

RESULT 14

HB2695 DNA-3-methyladenine glycosidase XF1326 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: HB2695
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: H82695
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-216 <SIM>
A;Cross-references: GB:AE003965; GB:AE003849; NID:99106313; PIDN:AAF84135.1; GSPDB:GN00
A;Experimental source: strain 9a5c
R;Simpon, A.J.G.; Reinach, F.C.; Aruda, P.; Abreu, F.A.; Aencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carrer, as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fro J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Lai chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marinho, C.L.; Marques, M.V.; Martins, C.; Accession: A38198
A;Authors: Martins, E.M.F.; Matsuoka, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y R;Seino, S.; Chen, L.; Seino, M.; Blondel, O.; Takeda, J.; Johnson, J.H.; Bell, G.I. Proc. Natl. Acad. Sci. U.S.A. 89, 584-588, 1992
A;Title: Cloning of the alpha 1 subunit of a voltage-dependent calcium channel expressed A;Reference number: A38198; MUID:92115705; PMID:1309948
A;Accession: A38198
A;Molecule type: mRNA
A;Residues: 1-2181 <SEI>

A;Contents: annotation

C;Genetics: XPI326
 A;Gene: XP1326
 Query Match 76.5%; Score 26; DB 2; Length 216;
 Best Local Similarity 71.4%; Pred. No. 44;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MMMMRRL 7
 Db 139 MMTLMLRL 145

RESULT 15
 H96811
 protein F3F9.20 [Imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Accession: H96811 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: H96811
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-302 <STO>
 A;Cross-references: GB:AE005173; NID:98052542; PIDN:AAF71806.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F3F9.20
 A;Map position: 1

Query Match 76.5%; Score 26; DB 2; Length 302;
 Best Local Similarity 71.4%; Pred. No. 60;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MMMMRRL 7
 Db 159 MMMMMMI 165

Search completed: February 11, 2004, 17:11:56
 Job time : 11.3333 secs

GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34%
Sequence: 1 MMMMMRL 7
Scoring table: BLUSUM62
Gapop 10.0 , Gapext 0.5
Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	79.4	10	1 US-09-417-279-2	Sequence 2, Appli
2	27	79.4	10	5 PCT-US94-14173-2	Sequence 2, Appli
3	27	79.4	61	4 US-09-252-991A-29802	Sequence 2, Appli
4	27	79.4	2161	1 US-07-745-206A-2	Sequence 2, Appli
5	27	79.4	2161	1 US-08-455-543A-49	Sequence 49, Appli
6	27	79.4	2161	1 US-08-455-543A-51	Sequence 51, Appli
7	27	79.4	2161	2 US-08-223-305C-49	Sequence 51, Appli
8	27	79.4	2161	2 US-08-223-305C-51	Sequence 51, Appli
9	27	79.4	2161	2 US-08-311-363-2	Sequence 2, Appli
10	26	76.5	385	1 US-08-891-254-3	Sequence 3, Appli
11	26	76.5	385	2 US-08-819-539-3	Sequence 3, Appli
12	26	76.5	385	5 PCT-US93-06243-2	Sequence 2, Appli
13	26	76.5	385	5 PCT-US96-08819-3	Sequence 3, Appli
14	26	76.5	403	2 US-08-200-724A-2	Sequence 3, Appli
15	26	76.5	403	2 US-09-030-270A-3	Sequence 3, Appli
16	26	76.5	403	3 US-08-851-376A-2	Sequence 2, Appli
17	26	76.5	403	3 US-08-984-207-3	Sequence 3, Appli
18	26	76.5	403	3 US-09-013-587-3	Sequence 3, Appli
19	26	76.5	403	4 US-09-086-118-23	Sequence 23, Appli
20	26	76.5	989	2 US-08-070-301-16	Sequence 16, Appli
21	25	73.5	62	1 US-08-488-961-6	Sequence 6, Appli
22	25	73.5	62	3 US-08-973-297-6	Sequence 6, Appli
23	25	73.5	62	5 PCT-US96-06511-6	Sequence 98, Appli
24	25	73.5	67	3 US-09-120-365-98	Sequence 98, Appli
25	25	73.5	67	3 US-09-515-039-98	Sequence 6508, Appli
26	25	73.5	185	4 US-09-328-352-6508	Sequence 4466, Appli
27	25	73.5	284	4 US-09-328-352-4466	Sequence 4466, Appli

ALIGNMENTS

RESULT 1
US-09-417-279-2
; Sequence 2, Application US/08417279
; Patent No. 5543498
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; APPLICANT: Igarashi, Michihiro
; TITLE OF INVENTION: Peptides to Overcome Inhibition of Nerve
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,279
; FILING DATE: 20005
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/166,350
; FILING DATE: 14-DEC-1993
; APPLICATION NUMBER: 08/162,480
; FILING DATE: 07-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3960001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; US-08-417-279-2

Query Match 5;保守性 Matches 5;保守性 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMMMMR 6

Db 1 MLLMMMR 6

RESULT 2
PCT-US94-14173-2
; Sequence 2, Application PC/TU9414173
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Peptides to Overcome Inhibition of Nerve
; TITLE OF INVENTION: Growth
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14173
; FILING DATE: 06-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/166,350
; FILING DATE: 14-DEC-1993
; CLASSIFICATION:
; PRIORITY NUMBER: 08/162,480
; FILING DATE: 07-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3960001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; PCT-US94-14173-2

Query Match 79.4%; Score 27; DB 5; Length 10;
Best Local Similarity 83.3%; Pred. No. 0.94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLLMMMR 6
Db 1 MLLMMMR 6

RESULT 3
US-09-252-991A-29802, Application US/09252991A
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenstein et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO: 29802
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-29802
Query Match 79.4%; Score 27; DB 4; Length 61;
Best Local Similarity 71.4%; Pred. No. 5.7;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLLMMMR 7
Db 3 MLLMMMR 9
RESULT 4
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and Methods
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. Lasalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/745,206A
; FILING DATE: 19910815
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2161 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-745-206A-2
Query Match 79.4%; Score 27; DB 1; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLLMMMR 6
Db 3 MLLMMMR 8
RESULT 5
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven

APPLICANT: Williams, Mark
 APPLICANT: Feldman, Daniel
 APPLICANT: McCue, Ann
 APPLICANT: Brenner, Robert
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brown, Martin, Haller & McClain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92101-2926
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/455,543A
 FILING DATE: May 31, 1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/223,305
 FILING DATE: April 4, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/745,206
 FILING DATE: 15-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,250
 FILING DATE: 30-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/482,384
 FILING DATE: 20-FEB-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/603,751
 FILING DATE: 04-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US89/01408
 FILING DATE: 04-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/176,899
 FILING DATE: 04-APR-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 6362-52517
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619)238-0999
 TELEFAX: (619)238-0062
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2161 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-455-543A-49

Query Match 1. MNNNNR 6 Score 27; DB 1; Length 2161;
 Best Local Similarity 83.3%; Pred. No. 2e+02; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 1. MNNNNR 6 Score 27; DB 1; Length 2161;
 Best Local Similarity 83.3%; Pred. No. 2e+02; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1. MNNNNR 6
 |||||:
 3 MNNNNR 8

Qy 1 MAMMR 6 Best Local Similarity 83.3%; Pred. No. 2e+02; Matches 5; Conservative 1; Mismatches 0; Indexes 0; Gaps 0;
 Db 3 MAMMR 8

RESULT 7 US-08-223-305C-49
 Sequence 49, Application US/08223305C
 Patent No. 5851824

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
 APPLICANT: Ellis, Steven
 APPLICANT: Williams, Mark
 APPLICANT: Feldman, Daniel
 APPLICANT: McCue, Ann
 APPLICANT: Brenner, Robert

TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McClain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92101-2926

COMPUTER READABLE FORM:

COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/223,305C
 FILING DATE: April 4, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/868,354
 FILING DATE: April 10, 1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/745,206
 FILING DATE: 15-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,250
 FILING DATE: 30-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/482,384
 FILING DATE: 20-FEB-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US89/01408
 FILING DATE: 04-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/176,899
 FILING DATE: 04-APR-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 52516 (P519739)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619)238-0999
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2161 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal

RESULT 8 US-08-223-305C-51
 Sequence 51, Application US/08223305C
 Patent No. 5851824

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
 APPLICANT: Ellis, Steven
 APPLICANT: Williams, Mark
 APPLICANT: Feldman, Daniel
 APPLICANT: McCue, Ann
 APPLICANT: Brenner, Robert

TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McClain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92101-2926

COMPUTER READABLE FORM:

COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/223,305C
 FILING DATE: April 4, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/745,206
 FILING DATE: 15-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,250
 FILING DATE: 30-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US89/01408
 FILING DATE: 04-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/176,899
 FILING DATE: 04-APR-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 52516 (P519739)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619)238-0999
 INFORMATION FOR SEQ ID NO: 51:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2161 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal

Query Match 79.4%; Score 27; DB 2; Length 2161;

US-08-223-305C-51

Query Match 79.4%; Score 27; DB 2; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels

Qy	1	MMMMMR	6
	3	:	
Db	3	MMMMMRK	8

RESULT 9
US-08-311-363-2
Sequence 2, Application US/08311363
Patent No. 5876958

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert

TITLE OF INVENTION: Human Calcium Channel Composit
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/311,363
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-51506
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-311-363-2

Query Match 79.4%; Score 27; DB 2; Len 10
Best Local Similarity 83.3%; Pred. No. 2e+02; Len 0;
Matches 5; Conservative 1; Mismatches 0;

Qy	1	MNNNNMR 6
	2	: 3
Db	3	NNNNNNMK 8

RESULT 10
US-08-891-254-3
; Sequence 3, Application US/08891254
; Patent No. 5776889
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min

APPLICANT: Beer, Steven V.
 TITLE OF INVENTION: Hypersensitive Response
 TITLE OF INVENTION: Induced Resistance In Plants
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 STREET: Clinton Square, P.O. Box 1051
 CITY: Rochester
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 14603

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.3
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/891,254
 FILING DATE: 10-JUL-1997
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/475,775
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldman, Michael L.
 REGISTRATION NUMBER: 30,727
 REFERENCE/DOCKET NUMBER: 14603/10050
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716) 263-1304
 TELEFAX: (716) 263-1600
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 385 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-891-254-3

RESULT 11
 US-08-819-539-3
 ; Sequence 3, Application US/08819539
 ; Patent No. 5859324
 ; GENERAL INFORMATION:
 ; APPLICANT: Wei, Zhong-Min.
 ; APPLICANT: Beer, Steven V.
 ; TITLE OF INVENTION: Hypersensitive Response
 ; TITLE OF INVENTION: Induced Resistance In Plants
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 ; STREET: Clinton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.3
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/819,539
 FILING DATE: 17-MAR-1997

CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/475,775
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldman, Michael L.
 REGISTRATION NUMBER: 30,727
 REFERENCE/DOCKET NUMBER: 14603/10050
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716) 263-1304
 TELEFAX: (716) 263-1600
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 385 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-819-539-3

Query Match 76.5%; Score 26; DB 2; Length 385;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Indels 0; Gaps 0;

Qy 1 MMNNMRL 7
 Db 63 MMNNMSM 69

RESULT 12
 PCT-US93-06243-2
 ; Sequence 2, Application PC/TUS9306243
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhong-Min Wei, David W. Bauer, Steven V. Laby
 ; ATTORNEY: Beer, Alan Collmer, Sheng-Yang He, and Ron J. Laby
 ; TITLE OF INVENTION: Elicitor of the Hypersensitive Response in Plants
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Yahwak & Associates
 ; STREET: 25 Skytop Drive
 ; CITY: Trumbull
 ; STATE: Connecticut
 ; COUNTRY: USA
 ; ZIP: 06611
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: Microsoft Word 4.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/06243
 ; FILING DATE: 19930630
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 907,935
 ; FILING DATE: 01-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: George M. Yahwak
 ; REGISTRATION NUMBER: 26,824
 ; REFERENCE/DOCKET NUMBER: CRF D-1172
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (203)268-1951
 ; TELEFAX: (203)268-1951
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 385 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; PCT-US93-06243-2

Query Match 76.5%; Score 26; DB 5; Length 385;

Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Indels 0; Gaps 0;

Qy 1 MMNNMRL 7
 Db 63 MMNNMSM 69

RESULT 13
 PCT-US96-08819-3
 ; Sequence 3, Application PC/TUS9608819
 ; GENERAL INFORMATION:
 ; APPLICANT: Cornell Research Foundation, Inc.
 ; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
 ; TITLE OF INVENTION: RESISTANCE IN PLANTS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 ; STREET: Clinton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.3.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/08819
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/475,775
 ; FILING DATE: 07-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldman, Michael L.
 ; REGISTRATION NUMBER: 30,727
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (716) 263-1304
 ; TELEFAX: (716) 263-1600
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 385 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US96-08819-3

Query Match 76.5%; Score 26; DB 5; Length 385;

Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Indels 0; Gaps 0;

Qy 1 MMNNMRL 7
 Db 63 MMNNMSM 69

RESULT 14
 US-08-200-724A-2
 ; Sequence 2, Application US/08200724A
 ; Patent No. 5849868
 ; GENERAL INFORMATION:
 ; APPLICANT: Wei, Zhong-Min
 ; APPLICANT: Bauer, David W.
 ; APPLICANT: Beer, Steven V.
 ; APPLICANT: Collmer, Alan
 ; APPLICANT: He, Sheng-Yang
 ; APPLICANT: Laby, Ron J.
 ; TITLE OF INVENTION: ELICITOR OF THE HYPERSENSITIVE RESPONSE
 ; TITLE OF INVENTION: IN PLANTS

Query Match 76.5%; Score 26; DB 5; Length 385;

NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 STREET: Clinton Square
 CITY: Rochester
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 14603
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/200,724A
 FILING DATE: 23-FEB-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldman, Michael L.
 REGISTRATION NUMBER: 30,727
 REFERENCE/DOCKET NUMBER: 19603/10030
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716) 263-1304
 TELEFAX: (716) 263-1600
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 403 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-200-724A-2

Query Match 76.5%; Score 26; DB 2; Length 403;

Best Local Similarity 71.4%; Pred. No. 64;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMMMRLL 7
 Db 63 MMMMSM 69

RESULT 15
 US-09-030-270A-3
 ; Sequence 3, Application US/09030270A
 ; Patent No. 5977060
 ; GENERAL INFORMATION:
 ; APPLICANT: Zitter, Thomas A.
 ; APPLICANT: Wei, Zhong-Min
 ; TITLE OF INVENTION: INSECT CONTROL WITH A
 ; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 ; STREET: P.O. Box 1051, Clinton Square
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/030,270A
 FILING DATE:
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/039,226
 FILING DATE: 28-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727
 REFERENCE/DOCKET NUMBER: 19603/1521
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716) 263-1304
 TELEFAX: (716) 263-1600
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 403 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-030-270A-3

Query Match 76.5%; Score 26; DB 2; Length 403;

Best Local Similarity 71.4%; Pred. No. 64;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMMMRLL 7

Db 63 MMMMSM 69

Search completed: February 11, 2004, 17:13:40
 Job time : 11.4167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 MMMMMVRL 7
Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5
Searched: 801455 seqs, 209382283 residues
Total number of hits satisfying chosen parameters: 801455
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	DB	ID	Description
1	34	100.0	7	10	US-09-901-187B-12	Sequence 12, Appl
2	32	94.1	103	9	US-09-864-761-44210	Sequence 44210, A
3	29	85.3	42	9	US-09-864-761-45051	Sequence 45051, A
4	28	82.4	32	9	US-09-864-761-34021	Sequence 34021, A
5	28	82.4	223	12	US-10-108-260A-4500	Sequence 4500, AP
6	28	82.4	255	12	US-10-017-161-2180	Sequence 2180, AP
7	28	82.4	255	12	US-10-292-798-1826	Sequence 1826, AP
8	27	79.4	18	9	US-09-864-761-35316	Sequence 35316, A
9	27	79.4	18	9	US-09-864-761-38185	Sequence 38185, A
10	27	79.4	33	9	US-09-864-761-45071	Sequence 45071, A
11	27	79.4	111	9	US-09-864-761-43445	Sequence 43445, A
12	27	79.4	305	10	US-09-886-055-35	Sequence 35, Appl
13	27	79.4	305	11	US-09-779-679-14	Sequence 14, Appl
14	27	79.4	305	11	US-09-779-679-16	Sequence 16, Appl
15	27	79.4	305	11	US-09-779-679-18	Sequence 18, Appl

RESULT 1
US-09-901-187B-12
; Sequence 12, Application US/09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Michael S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment of Invention: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-12

Query Match Similarity 100.0%; Score 34; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNNNNRL 7
Db 1 MNNNNRL 7
RESULT 2
US-09-864-761-44210
; Sequence 44210, Application US/09864761

Patent No. US20020048763A1
 GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aeonica-X-1
 CURRENT APPLICATION NUMBER: US 09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annotrax Sequence Listing Engine vers. 1.1
 SEQ ID NO 44210
 LENGTH: 103
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC005668.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 0.55
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
 US-09-864-761-44210
 Query Match 94.1%; Score 32; DB 9; Length 103;
 Best Local Similarity 85.7%; Pred. No. 3.6; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 M M M M M R L 7
 Db 16 M M M M M R M 22
 RESULT 3

US-09-864-761-45051
 Sequence 45051, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aeonica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annotrax Sequence Listing Engine vers. 1.1
 SEQ ID NO 45051
 LENGTH: 42
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC009933.5
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.67
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.69
 US-09-864-761-45051
 Query Match 85.3%; Score 29; DB 9; Length 42;
 Best Local Similarity 71.4%; Pred. No. 7.1;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 M M M M M R L 7
 Db 20 M M M M M K M 26

RESULT 4

US-09-864-761-34021 ; Sequence 34021, Application US/09864761
 Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 34021

LENGTH: 32

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC005100.2

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.77

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.74

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.76

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.81

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.78

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6

US-09-864-761-34021

Query Match 82.4%; Score 28; DB 9; Length 32;
 Best Local Similarity 85.7%; Pred. No. 9.2;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
 FILE REFERENCE: 084335/166
 CURRENT APPLICATION NUMBER: US/10/292,798
 CURRENT FILING DATE: 2002-11-13
 PRIOR APPLICATION NUMBER: 10/017,161
 PRIOR FILING DATE: 2001-12-18
 PRIOR APPLICATION NUMBER: JP 2001-246789
 PRIOR FILING DATE: 2001-06-18
 NUMBER OF SEQ ID NOS: 2070
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 1826
 LENGTH: 255
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-292-798-1826

Query Match 82.4%; Score 28; DB 12; Length 255;
 Best Local Similarity 71.4%; Pred. No. 68;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPMNMRL 7
 Db 111 MPMNMR 117

RESULT 8
 US-09-864-761-35316
 Sequence 35316, Application US/09864761
 Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
 SEQ ID NO 35316
 LENGTH: 18
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AL035704.7
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
 US-09-864-761-35316
 Query Match 79.4%; Score 27; DB 9; Length 18;
 Best Local Similarity 83.3%; Pred. No. 8.8;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPMNMR 6
 Db 13 MPMNMR 18

RESULT 9
 US-09-864-761-38185
 Sequence 38185, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aeomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30

PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 38185
 LENGTH: 18
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AL035704.9
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.8
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.1
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
 US-09-864-761-38185

Query Match 79.4%; Score 27; DB 9; Length 33;
 Best Local Similarity 83.3%; Pred. No. 16;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1 MNNNNMR 6	1 MNNNNMR 6
	13 MNNNNMRK 18	10 MNNNNMRK 15
Db		

RESULT 10
 US-09-864-761-45071
 Sequence 45071, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aeomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR APPLICATION NUMBER: PCT/US01/00663

RESULT 11
 US-09-864-761-43445
 Sequence 43445, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aeomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/2334,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: AnnoMAX Sequence Listing Engine vers. 1.1
 SEQ ID NO: 43445
 LENGTH: 111
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC006097.1
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.62
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.81
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
 OTHER INFORMATION: SWISSPROT HIT: Q10699, EVALUE 5.90e-01
 US-09-864-761-43445

Query Match 79.4%; Score 27; DB 9; Length 111;
 Best Local Similarity 83.3%; Pred. No. 51;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAMMAMR 6
 Db 1 MAMMAMK 6

RESULT 12
 US-09-886-055-35
 Sequence 35, Application US/09886055
 Patent No. US20020132273A1
 GENERAL INFORMATION:
 APPLICANT: STRYER, LUBERT
 TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
 FILE REFERENCE: 078003-0277150
 CURRENT FILING DATE: 2001-06-22
 PRIOR APPLICATION NUMBER: 60/213,812
 PRIOR FILING DATE: 2000-06-22
 NUMBER OF SEQ ID NOS: 522
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 35
 LENGTH: 305
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-886-055-35

Query Match 79.4%; Score 27; DB 10; Length 305;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAMMAMR 7
 Db 45 MINLMRL 51

RESULT 13
 US-09-779-679-14

Query Match 79.4%; Score 27; DB 11; Length 305;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAMMAMR 7
 Db 45 MINLMRL 51

RESULT 14
 US-09-779-679-16

Sequence 16, Application US/09779679
 Publication No. US20030082757A1
 GENERAL INFORMATION:
 APPLICANT: Taupier, Raymond J
 APPLICANT: Burgess, Catherine E
 APPLICANT: Padigaru, Muralidhara
 APPLICANT: Tchernev, Velizar T
 APPLICANT: Mishra, Vishnu S
 APPLICANT: Casman, Stacie
 APPLICANT: Ballinger, Robert
 APPLICANT: Vernet, Corine A
 APPLICANT: Li, Li
 APPLICANT: Spytek, Kimberly A
 APPLICANT: Andrew, David P
 APPLICANT: Mezes, Peter S

TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Same
FILE REFERENCE: 15966-661
CURRENT APPLICATION NUMBER: US/09/779,679
CURRENT FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: USSN 60/181045
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: USSN 60/183191
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: USSN 60/180,929
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: USSN 60/219758
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: USSN 60/181339
PRIOR FILING DATE: 2000-02-09
PRIOR APPLICATION NUMBER: USSN 60/181344
PRIOR FILING DATE: 2000-02-09
PRIOR APPLICATION NUMBER: USSN 60/221341
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: USSN 60/181392
PRIOR FILING DATE: 2000-02-09
PRIOR APPLICATION NUMBER: USSN 60/219585
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: USSN 60/181157
PRIOR FILING DATE: 2000-02-09
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 16
LENGTH: 305
TYPE: PRT
ORGANISM: Homo sapiens
US-09-779-679-16

Query Match 79.4%; Score 27; DB 11; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNNMRL 7
Db 45 MIMLMRL 51

Search completed: February 11, 2004, 17:54:14
Job time : 25.6667 secs

RESULT 15
US-09-779-679-18
Sequence 18, Application US/09779679
Publication No. US20030082757A1
GENERAL INFORMATION:
APPLICANT: Taupier, Raymond J
APPLICANT: Burgess, Catherine E
APPLICANT: Padigaru, Muralidhara
APPLICANT: Tchernev, Velizar T
APPLICANT: Mishra, Vishnu S
APPLICANT: Casman, Stacie
APPLICANT: Ballinger, Robert
APPLICANT: Vernet, Corine A
APPLICANT: Li, Li
APPLICANT: Spytek, Kimberly A
APPLICANT: Andrew, David P
APPLICANT: Mezes, Peter S
TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Same
FILE REFERENCE: 15966-661
CURRENT APPLICATION NUMBER: US/09/779,679
CURRENT FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: USSN 60/181045
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: USSN 60/183191
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: USSN 60/180,929
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: USSN 60/219758
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: USSN 60/181339
PRIOR FILING DATE: 2000-02-09
PRIOR APPLICATION NUMBER: USSN 60/181344

PT synuclein useful for treating a neurodegenerative disease involves
 PT determining aggregation of alpha synuclein in the presence of exogenous
 PT iron or copper -
 XX
 PS Claim 40; Page 37; 52pp; English.

CC The invention relates to screening of inhibitors of alpha-synuclein
 CC aggregation in the presence of exogenous iron or copper. The inhibitors
 CC are magnesium and alpha-synuclein binding peptides, which are
 CC useful for treating neurodegenerative disease that involves
 CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
 CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
 CC atrophy and Hallervorden-Spatz disease. The present sequence is a
 CC peptide that binds to the NAC (non-amyloid-beta protein component)
 CC portion of human alpha-synuclein and inhibits its aggregation.

XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 34; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 3
 Qy 1 MMMMMRL 7
 Db 1 MMMMMRL 7
 XX
 DT 04-FEB-2002 (first entry)
 XX
 ABG55058 standard; Peptide; 103 AA.
 XX
 AC ABG55058;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID No 33706.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 XX
 PS Claim 27; SEQ ID No 33706; 658pp; English.
 XX
 The invention relates to a single exon nucleic acid probe (SENTP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 103 AA;
 Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMRL 7
 Db 16 MMMMMRM 22

XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #7431 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PS Claim 27; SEQ ID NO 32560; 639pp + sequence listing; English.
 XX
 The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 103 AA;
 Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;

PR 04-FEB-2000; 20000US-0180312.
 PR 26-MAY-2000; 20000US-0207456.
 PR 30-JUN-2000; 20000US-0608408.
 PR 03-AUG-2000; 20000US-0632366.
 PR 21-SEP-2000; 20000US-0234687.
 PR 27-SEP-2000; 20000US-0236359.
 PR 04-OCT-2000; 20000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX PS Claim 27; SEQ ID No 33813; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENPs;
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX SQ Sequence 103 AA;
 XX Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX QY 1 M M M M R L 7
 16 M M M M R M 22
 DB 19-AUG-2002 (First entry)
 XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 32860.
 XX AC ABG43195;
 XX DT 19-AUG-2002 (First entry)
 XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 32860.
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD; chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; primary alveolar proteinosis; pulmonary hypertension;
 KW primary ciliary dyskinesia; pulmonary dysplasia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX OS Homo sapiens.
 XX PN WO200186003-A2.
 XX PD 15-NOV-2001.
 XX PF 30-JAN-2001; 2001WO-US00665.
 XX PR 04-FEB-2000; 20000US-180312P.
 PR 26-MAY-2000; 20000US-207456P.
 PR 30-JUN-2000; 20000US-0608408.
 PR 03-AUG-2000; 20000US-0632366.
 PR 21-SEP-2000; 20000US-234687P.
 PR 27-SEP-2000; 20000US-236359P.
 PR 04-OCT-2000; 20000GB-0024263.
 XX PA XX
 XX PI XX
 XX DR XX
 XX WPI; 2002-114183/15.
 XX PT XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX XX
 XX PS Claim 27; SEQ ID No 32860; 634pp; English.
 XX CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC their complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes having high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA, and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, lymphangioleiomyomatosis,
 CC haemosiderosis, pulmonary histiocytosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 103 AA;
 XX Query Match 94.1%; Score 32; DB 23; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX QY 1 M M M M R L 7
 XX Db 16 M M M M R M 22
 XX RESULT 8
 XX ID AAG24998 standard; Protein; 167 AA.
 XX AC AAG24998;
 XX DT 17-OCT-2000 (First entry)
 XX XX Arabidopsis thaliana protein fragment SEQ ID NO: 28884.
 XX DE XX

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

XX	XX	25-FEB-2000 ; 2000EP-0301439.	PR	23-JUN-1999;
PN	EP1033405-A2.	PR	24-JUN-1999;	
XX	06-SEP-2000.	PR	28-JUN-1999;	
PD		PR	29-JUN-1999;	
XX		PR	30-JUN-1999;	
PF		PR	01-JUL-1999;	
XX		PR	02-JUL-1999;	
PR	25-FEB-1999;	99US-0121825.	PR	06-JUL-1999;
PR	05-MAR-1999;	99US-0123180.	PR	07-JUL-1999;
PR	09-MAR-1999;	99US-0123548.	PR	08-JUL-1999;
PR	23-MAR-1999;	99US-0125788.	PR	09-JUL-1999;
PR	25-MAR-1999;	99US-0126264.	PR	10-JUL-1999;
PR	29-MAR-1999;	99US-0126785.	PR	11-JUL-1999;
PR	01-APR-1999;	99US-0127462.	PR	12-JUL-1999;
PR	06-APR-1999;	99US-0128234.	PR	13-JUL-1999;
PR	08-APR-1999;	99US-0128714.	PR	14-JUL-1999;
PR	16-APR-1999;	99US-0129845.	PR	15-JUL-1999;
PR	19-APR-1999;	99US-0130077.	PR	16-JUL-1999;
PR	21-APR-1999;	99US-0130449.	PR	17-JUL-1999;
PR	23-APR-1999;	99US-0130510.	PR	18-JUL-1999;
PR	23-APR-1999;	99US-0130891.	PR	19-JUL-1999;
PR	28-APR-1999;	99US-0131449.	PR	20-JUL-1999;
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;
PR	04-MAY-1999;	99US-0132407.	PR	22-JUL-1999;
PR	05-MAY-1999;	99US-0132484.	PR	23-JUL-1999;
PR	06-MAY-1999;	99US-0132485.	PR	24-JUL-1999;
PR	06-MAY-1999;	99US-0132486.	PR	25-JUL-1999;
PR	07-MAY-1999;	99US-0132487.	PR	26-JUL-1999;
PR	11-MAY-1999;	99US-0132863.	PR	27-JUL-1999;
PR	14-MAY-1999;	99US-0134256.	PR	28-JUL-1999;
PR	14-MAY-1999;	99US-0134218.	PR	29-JUL-1999;
PR	14-MAY-1999;	99US-0134219.	PR	30-JUL-1999;
PR	14-MAY-1999;	99US-0134221.	PR	31-JUL-1999;
PR	14-MAY-1999;	99US-0134370.	PR	01-AUG-1999;
PR	18-MAY-1999;	99US-0134768.	PR	02-AUG-1999;
PR	19-MAY-1999;	99US-0134941.	PR	03-AUG-1999;
PR	20-MAY-1999;	99US-0135124.	PR	04-AUG-1999;
PR	21-MAY-1999;	99US-0135353.	PR	05-AUG-1999;
PR	24-MAY-1999;	99US-0135629.	PR	06-AUG-1999;
PR	25-MAY-1999;	99US-0136021.	PR	07-AUG-1999;
PR	27-MAY-1999;	99US-0136392.	PR	08-AUG-1999;
PR	28-MAY-1999;	99US-0136782.	PR	09-AUG-1999;
PR	01-JUN-1999;	99US-0137222.	PR	10-AUG-1999;
PR	03-JUN-1999;	99US-0137528.	PR	11-AUG-1999;
PR	04-JUN-1999;	99US-0137502.	PR	12-AUG-1999;
PR	07-JUN-1999;	99US-0137724.	PR	13-AUG-1999;
PR	08-JUN-1999;	99US-0138094.	PR	14-AUG-1999;
PR	10-JUN-1999;	99US-0138540.	PR	15-AUG-1999;
PR	10-JUN-1999;	99US-0138847.	PR	16-AUG-1999;
PR	14-JUN-1999;	99US-0139119.	PR	17-AUG-1999;
PR	16-JUN-1999;	99US-0139452.	PR	18-AUG-1999;
PR	18-JUN-1999;	99US-0139453.	PR	19-AUG-1999;
PR	18-JUN-1999;	99US-0139492.	PR	20-AUG-1999;
PR	18-JUN-1999;	99US-0139454.	PR	21-AUG-1999;
PR	18-JUN-1999;	99US-0139455.	PR	22-AUG-1999;
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;
PR	18-JUN-1999;	99US-0139461.	PR	24-AUG-1999;
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;
PR	18-JUN-1999;	99US-0139470.	PR	27-AUG-1999;
PR	21-JUN-1999;	99US-0139817.	PR	28-AUG-1999;
PR	22-JUN-1999;	99US-0139899.	PR	29-AUG-1999;
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;

PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 XX PR 05-MAR-1999; 99US-0123180.
 XX PR 09-MAR-1999; 99US-0123548.
 XX PR 23-MAR-1999; 99US-0125788.
 XX PR 25-MAR-1999; 99US-0126264.
 XX PR 29-MAR-1999; 99US-0126785.
 XX PR 01-APR-1999; 99US-0127462.
 XX PR 06-APR-1999; 99US-0128234.
 XX PR 08-APR-1999; 99US-0128714.
 XX PR 16-APR-1999; 99US-0129845.
 XX PR 19-APR-1999; 99US-0130077.
 XX PR 21-APR-1999; 99US-0130449.
 XX PR 23-APR-1999; 99US-0130510.
 XX PR 23-APR-1999; 99US-0130891.
 XX PR 28-APR-1999; 99US-0131449.
 XX PR 30-APR-1999; 99US-0132048.
 XX PR 30-APR-1999; 99US-0132407.
 XX PR 04-MAY-1999; 99US-0132484.
 XX PR 05-MAY-1999; 99US-0132485.
 XX PR 06-MAY-1999; 99US-0132486.
 XX PR 06-MAY-1999; 99US-0132487.
 XX PR 07-MAY-1999; 99US-0132863.
 XX PR 07-MAY-1999; 99US-0132407.
 XX PR 11-MAY-1999; 99US-0134256.
 XX PR 14-MAY-1999; 99US-0134218.
 XX PR 14-MAY-1999; 99US-0134219.
 XX PR 14-MAY-1999; 99US-0134221.
 XX PR 14-MAY-1999; 99US-0134370.
 XX PR 18-MAY-1999; 99US-0134768.
 XX PR 19-MAY-1999; 99US-0134941.
 XX PR 20-MAY-1999; 99US-0135124.
 XX PR 21-MAY-1999; 99US-0135353.
 XX PR 24-MAY-1999; 99US-0135629.
 XX PR 25-MAY-1999; 99US-0136021.
 XX PR 27-MAY-1999; 99US-0136392.
 XX PR 28-MAY-1999; 99US-0136782.
 XX PR 01-JUN-1999; 99US-0137222.
 XX PR 03-JUN-1999; 99US-0137528.
 XX PR 04-JUN-1999; 99US-0137502.
 XX PR 07-JUN-1999; 99US-0137724.
 XX PR 08-JUN-1999; 99US-0138094.
 XX PR 10-JUN-1999; 99US-0138540.
 XX PR 10-JUN-1999; 99US-0138847.
 XX PR 14-JUN-1999; 99US-0139119.
 XX PR 16-JUN-1999; 99US-0139452.
 XX PR 16-JUN-1999; 99US-0139453.
 XX PR 17-JUN-1999; 99US-0139492.
 XX PR 18-JUN-1999; 99US-0139454.
 XX PR 18-JUN-1999; 99US-0139455.
 XX PR 18-JUN-1999; 99US-0139456.
 XX PR 18-JUN-1999; 99US-0139457.
 XX PR 18-JUN-1999; 99US-0139458.
 XX PR 18-JUN-1999; 99US-0139463.
 XX PR 18-JUN-1999; 99US-0139750.
 XX PR 21-JUN-1999; 99US-0139763.
 XX PR 22-JUN-1999; 99US-0139899.
 XX PR 23-JUN-1999; 99US-0140353.
 XX PR 23-JUN-1999; 99US-0140354.
 XX PR 24-JUN-1999; 99US-0140695.
 XX PR 28-JUN-1999; 99US-0140823.
 XX PR 29-JUN-1999; 99US-0140991.
 XX PR 30-JUN-1999; 99US-0141287.
 XX PR 01-JUL-1999; 99US-0141842.
 XX PR 01-JUL-1999; 99US-0142154.

Query Match 88.2%; Score 30; DB 21; Length 167;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNMMR 6
 DB 92 MNNMMR 97

RESULT 9
 AAG51419 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 65257.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX EP1033405-A2.
 PN

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 07-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147302.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999;	99US-0155139
PR 23-SEP-1999;	99US-0155486
PR 24-SEP-1999;	99US-0155659
PR 28-SEP-1999;	99US-0156458
PR 29-SEP-1999;	99US-0156596
PR 04-OCT-1999;	99US-0157117
PR 05-OCT-1999;	99US-0157753
PR 06-OCT-1999;	99US-0157865
PR 07-OCT-1999;	99US-0158029
PR 08-OCT-1999;	99US-0158232
PR 12-OCT-1999;	99US-0158369
PR 13-OCT-1999;	99US-0159293
PR 13-OCT-1999;	99US-0159294
PR 13-OCT-1999;	99US-0159295
PR 14-OCT-1999;	99US-0159329
PR 14-OCT-1999;	99US-0159330
PR 14-OCT-1999;	99US-0159331
PR 14-OCT-1999;	99US-0159637
PR 14-OCT-1999;	99US-0159638
PR 18-OCT-1999;	99US-0159584
PR 21-OCT-1999;	99US-0160741
PR 21-OCT-1999;	99US-0160767
PR 21-OCT-1999;	99US-0160768
PR 21-OCT-1999;	99US-0160770
PR 21-OCT-1999;	99US-0160814
PR 21-OCT-1999;	99US-0160815
PR 22-OCT-1999;	99US-0160980
PR 22-OCT-1999;	99US-0160981
PR 22-OCT-1999;	99US-0160989
PR 25-OCT-1999;	99US-0161404
PR 25-OCT-1999;	99US-0161405
PR 25-OCT-1999;	99US-0161406
PR 26-OCT-1999;	99US-0161359
PR 26-OCT-1999;	99US-0161360
PR 26-OCT-1999;	99US-0161361
PR 28-OCT-1999;	99US-0161920
PR 28-OCT-1999;	99US-0161992
PR 28-OCT-1999;	99US-0161993
PR 29-OCT-1999;	99US-0162142

PR	09-MAR-1999;	99US-0123548.	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	99US-0125788.	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	99US-0126264.	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	99US-0126785.	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	99US-0127462.	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	99US-0128234.	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	99US-0128714.	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	99US-0129845.	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	99US-0130077.	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	06-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	06-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	18-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	19-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	20-MAY-1999;	99US-0135124.	PR	28-JUL-1999;	99US-0145951.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146386.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146388.
PR	25-MAY-1999;	99US-0136021.	PR	02-AUG-1999;	99US-0146389.
PR	27-MAY-1999;	99US-0136392.	PR	03-AUG-1999;	99US-0147038.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147204.
PR	01-JUN-1999;	99US-0137222.	PR	04-AUG-1999;	99US-0147302.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147192.
PR	04-JUN-1999;	99US-0137502.	PR	05-AUG-1999;	99US-0147260.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147303.
PR	08-JUN-1999;	99US-0136782.	PR	06-AUG-1999;	99US-0147204.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138847.	PR	09-AUG-1999;	99US-0147935.
PR	14-JUN-1999;	99US-0139119.	PR	10-AUG-1999;	99US-0148171.
PR	16-JUN-1999;	99US-0139452.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139453.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	99US-0139454.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139455.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139460.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0154779.
PR	29-JUN-1999;	99US-0140991.	PR	22-SEP-1999;	99US-0155139.
PR	30-JUN-1999;	99US-0141287.	PR	23-SEP-1999;	99US-0155486.
PR	01-JUL-1999;	99US-0141842.	PR	24-SEP-1999;	99US-0155659.
PR	01-JUL-1999;	99US-0142154.	PR	28-SEP-1999;	99US-0156458.
PR	02-JUL-1999;	99US-0142055.	PR	29-SEP-1999;	99US-0156596.
PR	06-JUL-1999;	99US-0142390.	PR	04-OCT-1999;	99US-0157117.
PR	08-JUL-1999;	99US-0142803.	PR	05-OCT-1999;	99US-0157753.

PR 06-OCT-1999; 99US-0157865. PR 16-APR-1999; 99US-0129845.
 PR 07-OCT-1999; 99US-0158029. PR 19-APR-1999; 99US-0130077.
 PR 08-OCT-1999; 99US-0158232. PR 21-APR-1999; 99US-0130449.
 PR 12-CCT-1999; 99US-0158369. PR 23-APR-1999; 99US-0130510.
 PR 13-CCT-1999; 99US-0159293. PR 23-APR-1999; 99US-0130891.
 PR 13-OCT-1999; 99US-0159294. PR 28-APR-1999; 99US-0131449.
 PR 13-OCT-1999; 99US-0159295. PR 30-APR-1999; 99US-0132048.
 PR 14-OCT-1999; 99US-0159329. PR 30-APR-1999; 99US-0132407.
 PR 14-OCT-1999; 99US-0159330. PR 04-MAY-1999; 99US-0132484.
 PR 14-OCT-1999; 99US-0159331. PR 05-MAY-1999; 99US-0132485.
 PR 14-OCT-1999; 99US-0159637. PR 06-MAY-1999; 99US-0132486.
 PR 14-OCT-1999; 99US-0159638. PR 07-MAY-1999; 99US-0132487.
 PR 18-OCT-1999; 99US-0159584. PR 07-MAY-1999; 99US-0132863.
 PR 21-OCT-1999; 99US-0160741. PR 07-MAY-1999; 99US-0132863.
 PR 21-OCT-1999; 99US-0160767. PR 11-MAY-1999; 99US-0134256.
 PR 21-OCT-1999; 99US-0160768. PR 14-MAY-1999; 99US-0134218.
 PR 21-OCT-1999; 99US-0160770. PR 14-MAY-1999; 99US-0134219.
 PR 21-OCT-1999; 99US-0160814. PR 14-MAY-1999; 99US-0134370.
 PR 21-OCT-1999; 99US-0160815. PR 18-MAY-1999; 99US-0134768.
 PR 22-OCT-1999; 99US-0160980. PR 19-MAY-1999; 99US-0134941.
 PR 22-OCT-1999; 99US-0160981. PR 20-MAY-1999; 99US-0135124.
 PR 22-OCT-1999; 99US-0160989. PR 21-MAY-1999; 99US-0135353.
 PR 25-OCT-1999; 99US-0161404. PR 24-MAY-1999; 99US-0135629.
 PR 25-OCT-1999; 99US-0161405. PR 25-MAY-1999; 99US-0136021.
 PR 25-OCT-1999; 99US-0161406. PR 27-MAY-1999; 99US-0136392.
 PR 26-OCT-1999; 99US-0161359. PR 28-MAY-1999; 99US-0136782.
 PR 26-OCT-1999; 99US-0161360. PR 01-JUN-1999; 99US-0137222.
 PR 26-OCT-1999; 99US-0161361. PR 03-JUN-1999; 99US-0137528.
 PR 28-OCT-1999; 99US-0161920. PR 04-JUN-1999; 99US-0137502.
 PR 28-OCT-1999; 99US-0161992. PR 07-JUN-1999; 99US-0137724.
 PR 28-OCT-1999; 99US-0161993. PR 08-JUN-1999; 99US-0138094.
 PR 29-OCT-1999; 99US-0162142. PR 10-JUN-1999; 99US-0138540.
 PR 29-OCT-1999; 99US-0162142. PR 10-JUN-1999; 99US-0138847.
 PR 29-OCT-1999; 99US-0162142. PR 14-JUN-1999; 99US-0139119.
 PR 29-OCT-1999; 99US-0162142. PR 16-JUN-1999; 99US-0139452.
 PR 29-OCT-1999; 99US-0162142. PR 16-JUN-1999; 99US-0139453.
 PR 29-OCT-1999; 99US-0162142. PR 17-JUN-1999; 99US-0139492.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139454.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139455.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139456.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139457.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139493.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139458.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139459.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139460.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139461.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139457.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139462.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139463.
 PR 29-OCT-1999; 99US-0162142. PR 18-JUN-1999; 99US-0139750.
 PR 29-OCT-1999; 99US-0162142. PR 21-JUN-1999; 99US-0139763.
 PR 29-OCT-1999; 99US-0162142. PR 22-JUN-1999; 99US-0139817.
 PR 29-OCT-1999; 99US-0162142. PR 23-JUN-1999; 99US-0140353.
 PR 29-OCT-1999; 99US-0162142. PR 24-JUN-1999; 99US-0140354.
 PR 29-OCT-1999; 99US-0162142. PR 28-JUN-1999; 99US-0140823.
 PR 29-OCT-1999; 99US-0162142. PR 29-JUN-1999; 99US-0140991.
 PR 29-OCT-1999; 99US-0162142. PR 30-JUN-1999; 99US-0141287.
 PR 29-OCT-1999; 99US-0162142. PR 01-JUL-1999; 99US-0141842.
 PR 29-OCT-1999; 99US-0162142. PR 01-JUL-1999; 99US-0142154.
 PR 29-OCT-1999; 99US-0162142. PR 02-JUL-1999; 99US-0142055.
 PR 29-OCT-1999; 99US-0162142. PR 06-JUL-1999; 99US-0142390.
 PR 29-OCT-1999; 99US-0162142. PR 08-JUL-1999; 99US-0142803.
 PR 29-OCT-1999; 99US-0162142. PR 09-JUL-1999; 99US-0143624.
 PR 29-OCT-1999; 99US-0162142. PR 14-JUL-1999; 99US-0144005.
 PR 29-OCT-1999; 99US-0162142. PR 15-JUL-1999; 99US-0144085.
 PR 29-OCT-1999; 99US-0162142. PR 16-JUL-1999; 99US-0144086.
 PR 29-OCT-1999; 99US-0162142. PR 19-JUL-1999; 99US-0144325.
 PR 29-OCT-1999; 99US-0162142. PR 19-JUL-1999; 99US-0144332.
 PR 29-OCT-1999; 99US-0162142. PR 19-JUL-1999; 99US-0144333.

Query Match 88.2%; Score 30; DB 21; Length 194;
 Best Local Similarity 100.0%; Pred. No. 9.7; Mismatches 0; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 MAMMOR 6
Db	2 MAMMOR 7

RESULT 11
 AAG24997 DT 17-OCT-2000 (first entry)
 ID AAG24997 standard; Protein; 212 AA.
 XX AC AAG24997;
 XX DT 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 28883.
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28883.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PR 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.

PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 22-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 212;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy 1 MMMMMR 6
 Db 137 MMMMR 142

RESULT 12
 ID AAG51694 Standard; Protein; 212 AA.
 XX
 AC AAG51694;
 XX DT 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 65635.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 30-APR-1999; 99US-0132048.

PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	02-AUG-1999;	99US-0146389.
PR	19-MAY-1999;	99US-0134941.	PR	02-AUG-1999;	99US-0146389.
PR	20-MAY-1999;	99US-0135124.	PR	03-AUG-1999;	99US-0147038.
PR	21-MAY-1999;	99US-0135353.	PR	04-AUG-1999;	99US-0147204.
PR	24-MAY-1999;	99US-0135629.	PR	04-AUG-1999;	99US-0147302.
PR	25-MAY-1999;	99US-0136021.	PR	05-AUG-1999;	99US-0147192.
PR	27-MAY-1999;	99US-0136392.	PR	05-AUG-1999;	99US-0147260.
PR	28-MAY-1999;	99US-0136782.	PR	06-AUG-1999;	99US-0147303.
PR	01-JUN-1999;	99US-0137222.	PR	06-AUG-1999;	99US-0147416.
PR	03-JUN-1999;	99US-0137528.	PR	09-AUG-1999;	99US-0147493.
PR	04-JUN-1999;	99US-0137502.	PR	09-AUG-1999;	99US-0147935.
PR	07-JUN-1999;	99US-0137724.	PR	10-AUG-1999;	99US-0148171.
PR	08-JUN-1999;	99US-0138094.	PR	11-AUG-1999;	99US-0148319.
PR	10-JUN-1999;	99US-0138540.	PR	12-AUG-1999;	99US-0148341.
PR	10-JUN-1999;	99US-0138847.	PR	13-AUG-1999;	99US-0148565.
PR	14-JUN-1999;	99US-0139119.	PR	14-AUG-1999;	99US-0148684.
PR	16-JUN-1999;	99US-0139452.	PR	15-AUG-1999;	99US-0149368.
PR	16-JUN-1999;	99US-0139453.	PR	16-AUG-1999;	99US-0149175.
PR	17-JUN-1999;	99US-0139492.	PR	17-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139454.	PR	18-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139455.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139456.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139458.	PR	23-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139459.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139460.	PR	27-AUG-1999;	99US-0151080.
PR	18-JUN-1999;	99US-0139461.	PR	30-AUG-1999;	99US-0151303.
PR	18-JUN-1999;	99US-0139462.	PR	31-AUG-1999;	99US-0151438.
PR	18-JUN-1999;	99US-0139463.	PR	01-SEP-1999;	99US-0151930.
PR	18-JUN-1999;	99US-0139750.	PR	07-SEP-1999;	99US-0152363.
PR	18-JUN-1999;	99US-0139763.	PR	10-SEP-1999;	99US-0153070.
PR	21-JUN-1999;	99US-0139817.	PR	13-SEP-1999;	99US-0153758.
PR	22-JUN-1999;	99US-0139899.	PR	15-SEP-1999;	99US-0154018.
PR	23-JUN-1999;	99US-0140353.	PR	16-SEP-1999;	99US-0154039.
PR	23-JUN-1999;	99US-0140354.	PR	20-SEP-1999;	99US-0154779.
PR	01-JUL-1999;	99US-0141842.	PR	22-SEP-1999;	99US-0155139.
PR	01-JUL-1999;	99US-0142154.	PR	23-SEP-1999;	99US-0155486.
PR	02-JUL-1999;	99US-0142055.	PR	24-SEP-1999;	99US-0155659.
PR	06-JUL-1999;	99US-0142390.	PR	06-OCT-1999;	99US-0157865.
PR	08-JUL-1999;	99US-0142803.	PR	07-OCT-1999;	99US-0158029.
PR	09-JUL-1999;	99US-0140695.	PR	13-OCT-1999;	99US-0159295.
PR	12-JUL-1999;	99US-0140823.	PR	14-OCT-1999;	99US-0159329.
PR	13-JUL-1999;	99US-0140991.	PR	14-OCT-1999;	99US-0159330.
PR	30-JUN-1999;	99US-0141287.	PR	12-OCT-1999;	99US-0159331.
PR	01-JUL-1999;	99US-0144065.	PR	13-OCT-1999;	99US-0159294.
PR	02-JUL-1999;	99US-0142977.	PR	13-OCT-1999;	99US-0159637.
PR	13-JUL-1999;	99US-0143542.	PR	14-OCT-1999;	99US-0159638.
PR	14-JUL-1999;	99US-0144085.	PR	14-OCT-1999;	99US-0159584.
PR	15-JUL-1999;	99US-0144332.	PR	21-OCT-1999;	99US-0160741.
PR	16-JUL-1999;	99US-0144333.			
PR	19-JUL-1999;	99US-0144334.			
PR	19-JUL-1999;	99US-0144335.			
PR	20-JUL-1999;	99US-0144632.			
PR	20-JUL-1999;	99US-0144884.			
PR	21-JUL-1999;	99US-0144814.			
PR	21-JUL-1999;	99US-0145086.			

PR 21-OCT-1999; 99US-0160767. PR 14-MAY-1999; 99US-0134218.
 PR 21-OCT-1999; 99US-0160768. PR 14-MAY-1999; 99US-0134219.
 PR 21-OCT-1999; 99US-0160770. PR 14-MAY-1999; 99US-0134221.
 PR 21-OCT-1999; 99US-0160814. PR 14-MAY-1999; 99US-0134370.
 PR 21-OCT-1999; 99US-0160815. PR 18-MAY-1999; 99US-0134768.
 PR 22-OCT-1999; 99US-0160980. PR 19-MAY-1999; 99US-0134941.
 PR 22-OCT-1999; 99US-0160981. PR 20-MAY-1999; 99US-0135124.
 PR 22-OCT-1999; 99US-0160989. PR 21-MAY-1999; 99US-0135353.
 PR 25-OCT-1999; 99US-0161404. PR 24-MAY-1999; 99US-0135629.
 PR 25-OCT-1999; 99US-0161405. PR 25-MAY-1999; 99US-0136021.
 PR 25-OCT-1999; 99US-0161406. PR 27-MAY-1999; 99US-0136392.
 PR 26-OCT-1999; 99US-0161359. PR 28-MAY-1999; 99US-0136782.
 PR 26-OCT-1999; 99US-0161360. PR 01-JUN-1999; 99US-0137222.
 PR 26-OCT-1999; 99US-0161361. PR 03-JUN-1999; 99US-0137528.
 PR 28-OCT-1999; 99US-0161920. PR 04-JUN-1999; 99US-0137502.
 PR 28-OCT-1999; 99US-0161992. PR 07-JUN-1999; 99US-0137724.
 PR 28-OCT-1999; 99US-0161993. PR 08-JUN-1999; 99US-0138094.
 PR 29-OCT-1999; 99US-0162142. PR 10-JUN-1999; 99US-0138540.
 PR 29-OCT-1999; 99US-0162143. PR 10-JUN-1999; 99US-0138847.
 PR 29-OCT-1999; 99US-0162144. PR 14-JUN-1999; 99US-0139119.
 PR 29-OCT-1999; 99US-0162145. PR 16-JUN-1999; 99US-0139452.
 PR 29-OCT-1999; 99US-0162146. PR 16-JUN-1999; 99US-0139453.
 PR 29-OCT-1999; 99US-0162147. PR 17-JUN-1999; 99US-0139492.
 PR 29-OCT-1999; 99US-0162148. PR 18-JUN-1999; 99US-0139454.
 PR 29-OCT-1999; 99US-0162149. PR 18-JUN-1999; 99US-0139455.
 PR 29-OCT-1999; 99US-0162150. PR 18-JUN-1999; 99US-0139456.
 PR 29-OCT-1999; 99US-0162151. PR 18-JUN-1999; 99US-0139457.
 PR 29-OCT-1999; 99US-0162152. PR 18-JUN-1999; 99US-0139458.
 PR 29-OCT-1999; 99US-0162153. PR 18-JUN-1999; 99US-0139459.
 PR 29-OCT-1999; 99US-0162154. PR 18-JUN-1999; 99US-0139460.
 PR 29-OCT-1999; 99US-0162155. PR 18-JUN-1999; 99US-0139461.
 PR 29-OCT-1999; 99US-0162156. PR 18-JUN-1999; 99US-0139462.
 PR 29-OCT-1999; 99US-0162157. PR 18-JUN-1999; 99US-0139463.
 PR 29-OCT-1999; 99US-0162158. PR 18-JUN-1999; 99US-0139750.
 PR 29-OCT-1999; 99US-0162159. PR 18-JUN-1999; 99US-0139763.
 PR 29-OCT-1999; 99US-0162160. PR 18-JUN-1999; 99US-0139817.
 PR 29-OCT-1999; 99US-0162161. PR 18-JUN-1999; 99US-0139899.
 PR 29-OCT-1999; 99US-0162162. PR 18-JUN-1999; 99US-0140353.
 PR 29-OCT-1999; 99US-0162163. PR 18-JUN-1999; 99US-0140354.
 PR 29-OCT-1999; 99US-0162164. PR 18-JUN-1999; 99US-0140695.
 PR 29-OCT-1999; 99US-0162165. PR 18-JUN-1999; 99US-0140823.
 PR 29-OCT-1999; 99US-0162166. PR 18-JUN-1999; 99US-0140991.
 PR 29-OCT-1999; 99US-0162167. PR 18-JUN-1999; 99US-0141287.
 PR 29-OCT-1999; 99US-0162168. PR 01-JUL-1999; 99US-0141842.
 PR 29-OCT-1999; 99US-0162169. PR 01-JUL-1999; 99US-0142154.
 PR 29-OCT-1999; 99US-0162170. PR 02-JUL-1999; 99US-0142055.
 PR 29-OCT-1999; 99US-0162171. PR 06-JUL-1999; 99US-0142390.
 PR 29-OCT-1999; 99US-0162172. PR 08-JUL-1999; 99US-0142803.
 PR 29-OCT-1999; 99US-0162173. PR 09-JUL-1999; 99US-0142920.
 PR 29-OCT-1999; 99US-0162174. PR 12-JUL-1999; 99US-0142977.
 PR 29-OCT-1999; 99US-0162175. PR 13-JUL-1999; 99US-0143542.
 PR 29-OCT-1999; 99US-0162176. PR 14-JUL-1999; 99US-0143624.
 PR 29-OCT-1999; 99US-0162177. PR 15-JUL-1999; 99US-0144005.
 PR 29-OCT-1999; 99US-0162178. PR 16-JUL-1999; 99US-0144085.
 PR 29-OCT-1999; 99US-01626264. PR 16-JUL-1999; 99US-0144086.
 PR 29-OCT-1999; 99US-0126785. PR 19-JUL-1999; 99US-0144325.
 PR 01-APR-1999; 99US-0127462. PR 19-JUL-1999; 99US-0144331.
 PR 06-APR-1999; 99US-0128234. PR 19-JUL-1999; 99US-0144332.
 PR 08-APR-1999; 99US-0128714. PR 20-JUL-1999; 99US-0144334.
 PR 16-APR-1999; 99US-0129845. PR 20-JUL-1999; 99US-0144335.
 PR 19-APR-1999; 99US-0130077. PR 19-JUL-1999; 99US-0144336.
 PR 21-APR-1999; 99US-0130449. PR 21-JUL-1999; 99US-0144352.
 PR 23-APR-1999; 99US-0130510. PR 22-JUL-1999; 99US-0144353.
 PR 23-APR-1999; 99US-0130891. PR 22-JUL-1999; 99US-0145087.
 PR 28-APR-1999; 99US-0131449. PR 22-JUL-1999; 99US-0145089.
 PR 30-APR-1999; 99US-0132048. PR 21-JUL-1999; 99US-0145086.
 PR 30-APR-1999; 99US-0132407. PR 21-JUL-1999; 99US-0145088.
 PR 04-MAY-1999; 99US-0132484. PR 22-JUL-1999; 99US-0145085.
 PR 05-MAY-1999; 99US-0132485. PR 22-JUL-1999; 99US-0145087.
 PR 06-MAY-1999; 99US-0132486. PR 22-JUL-1999; 99US-0145192.
 PR 07-MAY-1999; 99US-0132487. PR 23-JUL-1999; 99US-0145145.
 PR 11-MAY-1999; 99US-0132863. PR 23-JUL-1999; 99US-0145218.

Query Match 88.2%; Score 30; DB 21; Length 212;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M~~MMMR~~ 6
 Db 137 M~~MMMR~~ 142

RESULT 13
 AAG23345 ID AAG23345 Standard; Protein: 214 AA.
 XX AC AAG23345;
 XX DT 17-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 26618.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 05-MAR-1999; 99US-0121825.
 PR 09-MAR-1999; 99US-0123180.
 PR 23-MAR-1999; 99US-0123548.
 PR 25-MAR-1999; 99US-0125788.
 PR 29-MAR-1999; 99US-0126264.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0134256.

PR 23-JUL-1999; 99US-0145224. PR 22-OCT-1999; 99US-0160989. PR 26-JUL-1999; 99US-0145276. PR 25-OCT-1999; 99US-0161404. PR 27-JUL-1999; 99US-0145913. PR 25-OCT-1999; 99US-0161406. PR 27-JUL-1999; 99US-0145919. PR 26-OCT-1999; 99US-0161359. PR 28-JUL-1999; 99US-0145951. PR 26-OCT-1999; 99US-0161360. PR 02-AUG-1999; 99US-0146386. PR 26-OCT-1999; 99US-0161361. PR 02-AUG-1999; 99US-0146388. PR 28-OCT-1999; 99US-0161920. PR 02-AUG-1999; 99US-0146389. PR 28-OCT-1999; 99US-0161921. PR 03-AUG-1999; 99US-0147038. PR 28-OCT-1999; 99US-0161993. PR 04-AUG-1999; 99US-0147204. PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 214; Best Local Similarity 100.0%; Pred. No. 11; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDDMMR 6
Db 2 MDDMMR 7

RESULT 14
ID AAG51436 standard; Protein; 233 AA.
XX AAG51436;
AC XX
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 65281.
DE XX
KW XX
Protein identification; Signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
KW XX
OS XX
Arabidopsis thaliana.
XX EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 04-MAY-1999; 99US-0132484.
PR 23-APR-1999; 99US-0130891.
PR 05-MAY-1999; 99US-0131449.
PR 06-MAY-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353. 99US-0146388.
 PR 24-MAY-1999; 99US-0135629. 99US-0146389.
 PR 25-MAY-1999; 99US-0136021. 99US-0147038.
 PR 27-MAY-1999; 99US-0136392. 99US-0147204.
 PR 28-MAY-1999; 99US-0136782. 99US-0147302.
 PR 01-JUN-1999; 99US-0137222. 99US-0147192.
 PR 03-JUN-1999; 99US-0137528. 99US-0147260.
 PR 04-JUN-1999; 99US-0137502. 99US-0147303.
 PR 07-JUN-1999; 99US-0137724. 99US-0147416.
 PR 08-JUN-1999; 99US-0138094. 99US-0147493.
 PR 10-JUN-1999; 99US-0138540. 99US-0147935.
 PR 10-JUN-1999; 99US-0138847. 99US-0148171.
 PR 14-JUN-1999; 99US-0139119. 99US-0148319.
 PR 16-JUN-1999; 99US-0139452. 99US-0148341.
 PR 18-JUN-1999; 99US-0139453. 99US-0148565.
 PR 17-JUN-1999; 99US-0139492. 99US-0148684.
 PR 18-JUN-1999; 99US-0139454. 99US-0149368.
 PR 18-JUN-1999; 99US-0139455. 99US-0149175.
 PR 18-JUN-1999; 99US-0139456. 99US-0149426.
 PR 18-JUN-1999; 99US-0139457. 99US-0149722.
 PR 18-JUN-1999; 99US-0139458. 99US-0149723.
 PR 18-JUN-1999; 99US-0139459. 99US-0149929.
 PR 18-JUN-1999; 99US-0139460. 99US-0149902.
 PR 18-JUN-1999; 99US-0139461. 99US-0149930.
 PR 18-JUN-1999; 99US-0139462. 99US-0150566.
 PR 18-JUN-1999; 99US-0139463. 99US-0150884.
 PR 18-JUN-1999; 99US-0139750. 99US-0151065.
 PR 18-JUN-1999; 99US-0139763. 99US-0151066.
 PR 21-JUN-1999; 99US-0139817. 99US-0151080.
 PR 22-JUN-1999; 99US-0139899. 99US-0151303.
 PR 23-JUN-1999; 99US-0140353. 99US-0151438.
 PR 23-JUN-1999; 99US-0140354. 99US-0151930.
 PR 24-JUN-1999; 99US-0140695. 99US-0152363.
 PR 28-JUN-1999; 99US-0140823. 99US-0153070.
 PR 29-JUN-1999; 99US-0140991. 99US-0155486.
 PR 30-JUN-1999; 99US-0141287. 99US-0155659.
 PR 01-JUL-1999; 99US-0141842. 99US-0156458.
 PR 01-JUL-1999; 99US-0142154. 99US-0156596.
 PR 02-JUL-1999; 99US-0142055. 99US-0157117.
 PR 06-JUL-1999; 99US-0142390. 99US-0157753.
 PR 08-JUL-1999; 99US-0142803. 99US-0158029.
 PR 09-JUL-1999; 99US-0142920. 99US-0158232.
 PR 12-JUL-1999; 99US-0142977. 99US-0158369.
 PR 13-JUL-1999; 99US-0143542. 99US-0159294.
 PR 14-JUL-1999; 99US-0143624. 99US-0159295.
 PR 15-JUL-1999; 99US-0144005. 99US-0159297.
 PR 16-JUL-1999; 99US-0144085. 99US-0159329.
 PR 16-JUL-1999; 99US-0144086. 99US-0159330.
 PR 19-JUL-1999; 99US-0144325. 99US-0159331.
 PR 19-JUL-1999; 99US-0144332. 99US-0159332.
 PR 19-JUL-1999; 99US-0144333. 99US-0159333.
 PR 19-JUL-1999; 99US-0144334. 99US-0159334.
 PR 20-JUL-1999; 99US-0144335. 99US-0159335.
 PR 20-JUL-1999; 99US-0144632. 99US-0159336.
 PR 20-JUL-1999; 99US-0144884. 99US-0159584.
 PR 21-JUL-1999; 99US-0144814. 99US-0160814.
 PR 21-JUL-1999; 99US-0145086. 99US-0160815.
 PR 21-JUL-1999; 99US-0145088. 99US-0160767.
 PR 22-JUL-1999; 99US-0145085. 99US-0160768.
 PR 22-JUL-1999; 99US-0145087. 99US-0160770.
 PR 22-JUL-1999; 99US-0145224. 99US-0160981.
 PR 26-JUL-1999; 99US-0145276. 99US-0161404.
 PR 27-JUL-1999; 99US-0145913. 99US-0161405.
 PR 27-JUL-1999; 99US-0145918. 99US-0161406.
 PR 27-JUL-1999; 99US-0145919. 99US-0161359.
 PR 28-JUL-1999; 99US-0145951. 99US-0161360.
 PR 02-AUG-1999; 99US-0146386. 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 233;
 Best Local Similarity 100.0%; Pred: No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M~~MM~~MR 6
 Db 2 M~~MM~~MR 7

RESULT 15
 AAG51693 ID AAG51693 standard; Protein; 235 AA.
 XX AC AAG51693;
 XX DT 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 65634.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 24-MAY-1999; 99US-0134768.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137229.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139464.
 PR 18-JUN-1999; 99US-0139465.
 PR 18-JUN-1999; 99US-0139466.
 PR 18-JUN-1999; 99US-0139467.
 PR 18-JUN-1999; 99US-0139468.
 PR 18-JUN-1999; 99US-0139469.
 PR 18-JUN-1999; 99US-0139470.
 PR 18-JUN-1999; 99US-0139471.
 PR 18-JUN-1999; 99US-0139472.
 PR 18-JUN-1999; 99US-0139473.
 PR 18-JUN-1999; 99US-0139474.
 PR 18-JUN-1999; 99US-0139475.
 PR 18-JUN-1999; 99US-0139476.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 20-JUL-1999; 99US-0144334.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMNNMR 6
 Db 160 MMNNMR 165

Search completed: February 11, 2004, 17:03:05
 Job time : 33.25 secs

PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156595.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159325.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match Similarity 88.2%; Score 30; DB 21; Length 235;
 Best Local Similarity 100.0%; Pred. No. 12;

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds

(without alignments)
 34.452 Million cell updates/sec

Title: US-09-901-187C-5
 Perfect score: 38
 Sequence: 1 THRLPSR 7

Scoring table: BLOSUM62

Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1993.DAT:*
- 14: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1994.DAT:*
- 15: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1995.DAT:*
- 16: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1996.DAT:*
- 17: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1997.DAT:*
- 18: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1998.DAT:*
- 19: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA1999.DAT:*
- 20: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA2000.DAT:*
- 21: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA2001.DAT:*
- 22: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA2002.DAT:*
- 23: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA2003.DAT:*
- 24: /SIDDS1/gcgdata/geneseq/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	38	100.0	7 20 AAY31365	B. cereus T spore
2	38	100.0	7 23 AAE14550	Human alpha-synuclein
3	38	100.0	1608 17 AAR98619	Borna disease virus
4	38	100.0	1711 17 AAR98605	Borna disease virus
5	35	92.1	163 22 AAU40892	Propionibacterium
6	34	89.5	356 22 ABG02462	Novel human diagno
7	33	86.8	9 16 AAR79677	Cyclin B/p33 (cdc2)
8	33	86.8	9 21 AAY57976	PP60-c-src peptide
9	33	86.8	18 22 ABG4096	Human liver peptid

Human peptide #724
 Peptide #753 encod
 Protein #709 encod
 Human brain expres
 Human bone marrow
 Peptide #734 encod
 Peptide #747 encod
 Peptide #711 encod
 Human peptide enco
 Human cardiovascular
 Propionibacterium
 Human secreted pro
 Human EST encoded
 Propionibacterium
 Mutant chicken c-S
 Chicken pp60 c-src
 Wild-type chicken
 Mutant chicken c-S
 Mutant chicken c-S
 Amino acid sequenc
 Arabidopsis thalia
 Propionibacterium
 Novel human diagno
 Human DNA/RNA bind
 Propionibacterium
 Propionibacterium
 Propionibacterium
 Novel human diagno
 Human hPPG-1 prote
 Propionibacterium
 Human RTA-like G p
 Human G-protein co
 Human G-protein co
 Human G protein-co
 Human G protein-co
 Human GPCR polypep
 Human mas proto-on

ALIGNMENTS

RESULT 1
 AAY31365
 ID AAY31365 standard; peptide; 7 AA.

XX 01-NOV-1999 (first entry)
 XX DE B. cereus T spore tight-binding peptide.
 XX XX B. cereus spore; biopanning; phage-display library; Bacillus; vaccine;
 XX KW Bacterial spore; biopanning; phage-display library; Bacillus; vaccine;
 XX KW pathogen detection; biological warfare agent; B. anthracis.
 XX XX Bacillus cereus.

XX XX WO9936081-A1.
 XX PD 22-JUL-1999.
 XX XX 99WO-US00771.
 XX PR 14-JAN-1998; 98US-0071411.
 XX XX (UABR-) UAB RES FOUND.
 PA PA Turnbough CL;
 XX PI DR N-PSDB; AAX99291.
 XX PT Identifying peptides that bind to the surface of bacterial spores by
 PT biopanning phage-display library, useful as vaccines and diagnostic

PT agents
 XX Disclosure; Page 13; 23pp; English.
 PS The invention provides peptides that bind to the surface of bacterial spores. These peptides are identified by biopanning a phage-display library with the spores. The library is incubated with spores and any phage-spore complexes formed recovered by centrifuging. They are washed thoroughly, then phage eluted with buffer, the eluate neutralised and eluted phage amplified. The procedure is repeated, for 3 or 4 rounds of panning, then individual clones purified, amplified and genomic DNA extracted for determination of peptide-encoding sequences. Peptides encoded by these sequences are then tested for ability to bind to target spores. The peptides are used for capture and identification of bacterial spores, particularly of the genus *Bacillus*, particularly for detecting pathogens, or otherwise harmful species, in the environment (e.g. air, water or food) or in clinical samples. They are also used to protect against disease-causing spores (e.g. by incorporation in protective masks), and in vaccines to generate a protective antibody response. The peptides-specific antibodies are useful as detection reagents, e.g. in enzyme-linked immunosorbent assay. The peptides are species specific, i.e. they can differentiate between the possible biological warfare agent B. anthracis and related species, widely present in the environment and likely to give rise to false positive results. Sequences AAY31365-372 represent amino acid sequences of *B. cereus* T spore tight-binding peptides.

XX Sequence 7 AA;
 PS Claim 40; Page 37; 52pp; English.
 XX The invention relates to screening of inhibitors of alpha-synuclein aggregation in the presence of exogenous iron or copper. The inhibitors are magnesium and alpha-synuclein-binding peptides, which are useful for treating neurodegenerative disease that involves the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system atrophy and Hallervorden-Spatz disease. The present sequence is a peptide that binds to C-terminal portion of human alpha-synuclein and inhibits its aggregation.

XX Sequence 7 AA;
 PS Query Match 100.0%; Score 38; DB 23; Length 7;
 PS Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 PS Matches 7; Conservative 0; Mismatches 0;
 PS Indels 0; Gaps 0;
 PS QY 1 THRLPSR 7
 PS DB 1 THRLPSR 7

RESULT 3
 AAR98619
 ID AAR98619 standard; Protein; 1608 AA.

XX Sequence 7 AA;
 PS Query Match 100.0%; Score 38; DB 20; Length 7;
 PS Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 PS Matches 7; Conservative 0; Mismatches 0;
 PS Indels 0; Gaps 0;
 PS QY 1 THRLPSR 7
 PS DB 1 THRLPSR 7

RESULT 2
 AAE14550
 ID AAE14550 standard; Peptide; 7 AA.

XX AC AAR98619;
 XX DT 10-DEC-1996 (First entry)
 XX XX Bornna disease virus polymerase.
 XX KW Bornna disease virus; BDV; G-protein; p57; nervous system disease;
 XX KW neuro-psychiatric disease; schizophrenia; diagnosis; therapy;
 XX KW vaccine; antibody.
 XX OS Bornna disease virus strain V.
 XX PN WO9621020-A2.
 XX PD 11-JUL-1996.
 XX PF 05-JAN-1996; 96WO-US00418.

XX Human alpha-synuclein aggregation inhibitor #5.
 XX Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
 XX Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
 XX multiple system atrophy; Hallervorden-Spatz disease; human.
 XX Homo sapiens.
 XX WO200204482-A1.
 XX PD 17-JAN-2002.

XX PF 06-JUL-2001; 2001WO-US21379.
 XX PR 07-JUL-2000; 2000US-217319P.
 XX PR 28-MAR-2001; 2001US-279199P.
 XX DR WPI; 2002-179695/23.

XX PA (PANA-) PANACEA PHARM INC.
 XX PI Wolozin B, Ostretsova-Golts N, Lebowitz MS;
 XX XX WPI; 2002-179695/23.
 XX Determination of an agent capable of inhibiting aggregation of alpha synuclein useful for treating a neurodegenerative disease involves determining aggregation of alpha synuclein in the presence of exogenous iron or copper -

XX CC Claim 40; Page 37; 52pp; English.
 XX CC The invention relates to screening of inhibitors of alpha-synuclein aggregation in the presence of exogenous iron or copper. The inhibitors are magnesium and alpha-synuclein-binding peptides, which are useful for treating neurodegenerative disease that involves the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system atrophy and Hallervorden-Spatz disease. The present sequence is a peptide that binds to C-terminal portion of human alpha-synuclein and inhibits its aggregation.

XX CC Sequence 7 AA;
 XX CC Query Match 100.0%; Score 38; DB 23; Length 7;
 XX CC Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 XX CC Matches 7; Conservative 0; Mismatches 0;
 XX CC Indels 0; Gaps 0;
 XX CC QY 1 THRLPSR 7
 XX CC DB 1 THRLPSR 7

XX CC RESULT 3
 XX CC AAR98619
 XX CC ID AAR98619 standard; Protein; 1608 AA.

XX CC Sequence 7 AA;
 XX CC AC AAR98619;
 XX CC DT 10-DEC-1996 (First entry)
 XX CC XX Bornna disease virus polymerase.
 XX CC KW Bornna disease virus; BDV; G-protein; p57; nervous system disease;
 XX CC KW neuro-psychiatric disease; schizophrenia; diagnosis; therapy;
 XX CC KW vaccine; antibody.
 XX CC OS Bornna disease virus strain V.
 XX CC PN WO9621020-A2.
 XX CC PD 11-JUL-1996.
 XX CC PF 05-JAN-1996; 96WO-US00418.

XX CC Human alpha-synuclein aggregation inhibitor #5.
 XX CC Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
 XX CC Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
 XX CC multiple system atrophy; Hallervorden-Spatz disease; human.
 XX CC Homo sapiens.
 XX CC WO200204482-A1.
 XX CC PD 17-JAN-2002.

XX CC PF 06-JUL-2001; 2001WO-US21379.
 XX CC PR 07-JUL-2000; 2000US-217319P.
 XX CC PR 28-MAR-2001; 2001US-279199P.
 XX CC DR WPI; 2002-179695/23.

XX CC PA (PANA-) PANACEA PHARM INC.
 XX CC PI Wolozin B, Ostretsova-Golts N, Lebowitz MS;
 XX CC XX WPI; 2002-179695/23.
 XX CC Determination of an agent capable of inhibiting aggregation of alpha synuclein useful for treating a neurodegenerative disease involves determining aggregation of alpha synuclein in the presence of exogenous iron or copper -

GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-5

Perfect score: 38

Sequence: 1 TIRLPSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:*

```

1: /cgn2_6/ptodata/4/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NNEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NNEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NNEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NNEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NNEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NNEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NNEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

```

Pre. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	38	100.0	7	10 US-09-901-187B-5	Sequence 5, App1
2	38	100.0	7	11 US-09-229-751A-72	Sequence 72, App1
3	33	86.8	18	9 US-09-864-761-34008	Sequence 34008, A
4	33	86.8	25	9 US-09-764-869-915	Sequence 915, App
5	33	86.8	25	12 US-10-227-577-915	Sequence 915, App
6	33	86.8	25	15 US-10-091-504-915	Sequence 915, App
7	33	86.8	362	12 US-10-369-493-10705	Sequence 10705, A
8	32	84.2	330	12 US-10-369-493-10663	Sequence 10663, A
9	32	84.2	542	12 US-10-369-493-12602	Sequence 12602, A
10	31	81.6	71	12 US-09-864-408A-5098	Sequence 5098, App
11	31	81.6	321	10 US-09-920-068A-2	Sequence 2, App1
12	31	81.6	321	12 US-10-017-161-1050	Sequence 1050, App
13	31	81.6	321	12 US-10-239-421-2	Sequence 2, App1
14	31	81.6	321	12 US-10-258-768-1	Sequence 1, App1
15	31	81.6	321	12 US-10-088-726-19	Sequence 19, App1

RESULT 1
US-09-901-187B-5
; Sequence 5, Application US 09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: WoLozin, Benjamin
; APPLICANT: Ostrzotova-Golts, Natalie
; APPLICANT: Lebowitz, Michael S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment of Invention: Alpha-Synuclein Diseases
; FILE REFERENCE: PANO1/002US
; CURRENT APPLICATION NUMBER: US 09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-5

Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 THRILPSR 7
DB 1 THRILPSR 7
RESULT 2
US-09-229-751A-72
; Sequence 72, Application US/09229751A

Publication No. US20030044838A1

GENERAL INFORMATION:
 APPLICANT: Turnbough, Charles K.
 TITLE OF INVENTION: PEPTIDE LIGANDS THAT BIND TO SURFACES
 OF BACTERIAL CELLS

NUMBER OF SEQUENCES: 80
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Glenna Hendricks
 STREET: P.O. Box 2509
 CITY: Fairfax
 STATE: VA
 COUNTRY: USA
 ZIP: 22031

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/229,751A
 FILING DATE: 14-Jan-1999

CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Hendricks, Glenna M.
 REGISTRATION NUMBER: 32,535

REFERENCE/DOCKET NUMBER: Turn
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 425-8405
 TELEFAX: (703) 425-8406

INFORMATION FOR SEQ ID NO: 72:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown

MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 72:

US-09-229-751A-72
 Query Match 100.0%; Score 38; DB 11; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FILE REFERENCE: Aeonica-X-1

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 SEQ ID NO 34008
 LENGTH: 18
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC006017.2
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.9
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN HUMAN HIT: AW997046.1, EVALUE 2.00e-04
 US-09-864-761-34008

Query Match 86.8%; Score 33; DB 9; Length 18;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
 US-09-764-869-915
 Sequence 915, Application US/09764869
 Patent No. US20020061521A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PC007
 CURRENT APPLICATION NUMBER: US/09/764,869
 CURRENT FILING DATE: 2001-01-17
 PRIOR application data removed - refer to PALM or file wrapper
 NUMBER OF SEQ ID NOS: 2442
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 915
 LENGTH: 25
 TYPE: PRT